UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

January 26, 2001

DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-DT-1, Revised

Memorandum For: Magdalena Ramos
Leader, ACE Implementation Team

From: Danny R. Childers DRC
Leader, Design Team

Subject: Accuracy and Coverage Evaluation: The Design Document

Attached is the documentation of the design of the Census 2000 Accuracy and Coverage Evaluation. If there are any questions, contact Danny R. Childers (301-457-4184).

cc. DSSD Census 2000 Procedures and Operations Memorandum Series Distribution List
ACE Implementation Team
Statistical Design Team Leaders
Chron

The Design of the Census 2000 Accuracy and Coverage Evaluation (ACE)

		Page
1.0	Introduction	1
2.0	The Housing Unit Phase	2
2.1	Independent Listing	2
2.2	Sampling	4
2.2.1	The Listing Sample	5
2.2.2	Medium and Large Block Cluster Reduction	6
2.2.3	Small Block Cluster Reduction	7
2.3	Obtain DMAF Extract	7
2.4	Before Follow-up Matching	8
2.4.1	Computer Match	8
2.4.2	Clerical Match	9
2.4.3	Technician Review	11
2.4.4	Analyst Review	12
2.5	Clusters With No Clerical Matching	12
2.6	Housing Unit Follow-up	13
2.7	After Follow-up Coding	15
2.7.1	Clerical Coding	15
2.7.2	Technician Review	16
2.7.3	Analyst Review	17
2.8	Relisting for Clusters with ACE Geocoding Errors	17
2.9	Reduction Within Large Block Clusters	18
2.9.1	Subsampling the P Sample and the E Sample	18
2.9.2	The Preliminary Enhanced List	20
2.9.3	The Subsampled Preliminary Enhanced List	21
2.9.4	The Enhanced List	21
3.0	The Interviewing Phase	22
3.1	Person Interview	22
3.1.1	Interviewing Outcome Codes	23
3.1.2	Edit for CAPI Data Review	24
3.1.3	ACE Mover and Residence Status Codes	25
3.2	ACE Form Selection	26
4.0	The Person Phase	27
4.1	P-Sample Identification	27
4.2	Preliminary P-sample Estimation Outcome Codes	27
4.3	E-Sample Identification	28
4.4	E-sample Insufficient Information for Matching and Follow-up	30
4.4.1	Census Data Defined	31
4.4.2	Computer Coding of Insufficient Information for Matching and Follow-up	32
4.4.3	Clerical Coding of Insufficient Information for Matching and	33

	Follow-up	
4.5	P-sample Insufficient Information for Matching and Follow-up	33
4.6	Within Block Cluster Matching	34
	4.6.1 Computer Matching	35
	4.6.2 Clerical Coding	36
	4.6.3 Census Images	39
	4.6.4 Duplicate Search Within Cluster	40
	4.6.5 Additional Clerical Coding	42
	4.6.5.1 Census Geocoding Errors	42
	4.6.5.2 Coding Nonmatches in Large Households	42
4.7	Targeted Extended Search	43
	4.7.1 P-sample Matching Extended Search	45
	4.7.2 E-sample Extended Search for Geocoding Errors	46
	4.7.3 E-sample Targeted Duplicate Search	47
	4.7.4 Added and Deleted Housing Units	48
4.8	Technician Review	49
4.9	Analyst Review	49
4.10	ACE Person Follow-up	49
	4.10.1 The ACE Person Follow-up Questionnaire	54
	4.10.2 ACE Person Follow-up Outcome Codes	55
4.11	After Follow-up Coding	56
4.12	Outlier Review	60
4.13	Additional Computer Coding	61
	4.13.1 Whole/Partial Code	61
	4.13.2 Address Code	62
	4.13.3 Basic Street Address Code	63
	4.13.4 Follow-up Flag	63
4.14	Final P-sample Person Match Codes	63
	4.14.1 Matched	63
	4.14.2 Not Matched	63
	4.14.3 Unresolved	64
	4.14.4 Removed from the P-sample	64
4.15	Final E-sample Person Enumeration Codes	65
	4.15.1 Correctly Enumerated	65
	4.15.2 Erroneously Enumerated	65
	4.15.3 Unresolved	66
4.16	Final P-sample Person Residence Codes	66
	4.16.1 Resident	67
	4.16.2 Nonresident	67
	4.16.3 Unresolved	68
4.17	Estimation Outcome Codes	68
	4.17.1 Final P-sample Estimation Outcome Codes for Census Day	68

The Design of the Census 2000 Accuracy and Coverage Evaluation

1.0 Introduction

The Census Bureau will conduct an Accuracy and Coverage Evaluation (ACE) to measure the overall and differential coverage of the U.S. population in Census 2000. The ACE will also provide base population figures for other Census Bureau programs, such as the Census Bureau's intercensal population estimates, American Community Survey, and other demographic surveys. Under the traditional census plan, the ACE will not be used to adjust the census figures for reapportionment purposes.

The major steps of the ACE are housing unit matching and person matching. Housing units within the sample clusters are listed and matched to the census inventory of housing units. After reconciling the nonmatches, a list of ACE housing units that are confirmed to have existed within the block clusters is prepared and person interviews are conducted for these P-sample housing units. Paper questionnaires are used for the housing unit listing and follow-up. The person interviewing is conducted by telephone and personal visit using Computer Assisted Personal Interview (CAPI).

Three types of people are collected in the person interview: those who lived at the sample address at the time of the interview and on census day (i.e., nonmovers), those who have moved into the sample address, since census day (i.e., inmovers), and those who lived at the sample address on census day but lived elsewhere at the time of the ACE interview (i.e., outmovers). Census day residence status is established within the CAPI instrument for the nonmovers and outmovers. Questions are not asked about other residences to establish residence status at the time of the person interview for the inmovers. We obtain a proxy interview for the outmovers, since they have moved. No outmover tracing is conducted for the 2000 ACE.

The "Population Sample" or "P-sample" and "Enumeration Sample" or "E-sample" have traditionally defined the samples for dual system estimation. The P-sample consists of people enumerated independent of the census. The E-sample consists of people enumerated in the census. After matching and reconciliation, the P-sample yields an estimate of the population missed in the census while the E-sample yields an estimate of the correctly enumerated people in the census. Putting these two components together yields an ACE estimate of census coverage.

There are non E-sample census people in large block clusters after subsampling. People counted in the census in institutional and noninstitutional group quarters are part of the non E-sample census people. Matching between the P-sample and non E-sample people is permitted, but non E-sample people who do not match are not sent for a follow-up interview. The correct enumeration rate is estimated from the E-sample. Matching between the P-sample and census people counted in group quarters is allowed in case there is misclassification of group quarters.

Each individual mobile home is assigned a unique map spot number whether the mobile home is in a park or not. The location of the mobile home is identified by placing the map spot number for the mobile home on the map. This is the same procedure that is used by the census.

The following items are collected and recorded in the listing book for each basic street address:

- City style addresses (house number and street name)
- Non-city style address (route numbers, route and box numbers, or any other type of address that is not a city style address)
- Householder name (rural areas only[2])
- Description of address (for non-house number addresses only in both urban and rural areas)
- Number of housing units in basic street address
- Type of basic street address (single unit, multi-unit, mobile home not in a mobile home park, mobile home in a mobile home park, housing unit in special place, multi-unit in a special place, or other)
- Unit status for single units (intended for occupancy, under construction, future construction, unfit for habitation, boarded up, storage of household goods, and other)

The following items are also collected and recorded in the listing book for each multi-unit within basic street address:

- unit description/apartment number
- Unit status for multi units (intended for occupancy, under construction, future construction, unfit for habitation, boarded up, storage of household goods, and other)

The following items are also collected and recorded in the listing book for each mobile home in a mobile home park:

- house number, lot number, or physical description
- street name
- rural address
- Unit status (intended for occupancy, unfit for habitation, boarded up, storage of household goods, vacant trailer site in a mobile home park, and other)

After the listing books are received in the National Processing Center (NPC) in Jeffersonville, they are checked in and keyed. The quality assurance is 100 percent for the keying of these listing books. Keying rejects are reviewed clerically to correct errors before the matching begins.

---

[2] We need the household name for matching when the ACE address is not house number and street name. We ask for the respondent name in all rural areas because the census address may not be house number and street name. We can not be guaranteed that the census address is collected in the same manner as the ACE.

2.2.1 The Listing Sample

The listing sample was selected in the second quarter of 1999. It was important to select the listing sample by June, 1999, in order to produce the materials required to start the listing operation by September, 1999. Under this schedule it was necessary to use early census address list information to select the ACE listing sample. The housing unit counts did not reflect all of the enhancements to the census address list which were incorporated by July, 1999 and did not include the final type of enumeration area classifications.

Under the ICM sampling plan, key features of the ACE listing sample selection include:

- roughly equal sample sizes for most states except the most populous
- a separate sample for American Indian Reservations
- roughly proportional allocation of sample within states
- a separate sample of small block clusters
- an oversample of large block clusters

The ACE primary sampling unit is the block cluster. A block cluster is a single census collection block or group of geographically contiguous census collection blocks. All census collection blocks in the 50 states, the District of Columbia and Puerto Rico are clustered except for collection blocks in remote Alaska areas. Remote Alaska is not in the sampling universe due to difficulties of the remote areas. The goal is to form block clusters which have identifiable boundaries and are of a manageable size in terms of both land area and number of housing units, roughly 30 housing units as indicated from the early census address list. Census blocks with 80 or more housing units were seldom clustered with other blocks except under specified situations. For ACE 2000, small census blocks were allowed to be clustered with a neighboring census block containing housing units. This clustering was limited in order to prevent the geographical size of the clusters from becoming too unwieldy. Also, an important feature of block clustering is to create clusters which have similar census address list creation methods, and thus similar census types of enumeration areas. This is important operationally so that blocks within a cluster have similar census materials and identification methods (e.g., map spots).

Using the early census address list housing unit counts, block clusters were stratified by size: small (0 to 2 housing units), medium (3 to 79 housing units) and large (80 or more housing units). In states with a sufficient number of American Indians living on reservations based on the 1990 census count, a separate sampling stratum was formed of American Indian Reservation block clusters. Within each sampling stratum, a systematic sample of block clusters was selected with equal probability. Note that there was a workload limitation within each state on the number of housing units that could be listed. If the expected number of housing units to list exceeded the workload limit, then the number of sample clusters from the large stratum was reduced to maintain the listing workload targets.

This stage of sampling yielded 29,136 block clusters and roughly 2 million housing units to be

2.2.3 Small Block Cluster Reduction

The small block cluster subsampling occurs in late January through early February 2000 following the completion of the keying of the independent listing books. It is important to wait for the results of the keying operation to get the best available number of housing units in the cluster. The small block cluster subsampling operation overlaps with the start of the housing unit matching. These two operations can overlap at this time because only the small block clusters selected in the ACE listing sample are eligible for this subsampling operation, and the housing unit matching can start with the other clusters. However, it is important to finish the small block cluster subsampling before the housing unit matching operation ends. Small block clusters must go to housing unit matching as well.

Before this operation the ACE reduction sample contains 5,000 small clusters in the United States and 96 small clusters in Puerto Rico. Small clusters are expected to have between zero and two housing units based on the early census address list. Conducting interviewing and follow-up operations in clusters of this size are not as cost-effective as in larger clusters. Therefore, to allocate ACE resources more efficiently, we will only include a subsample of these small clusters in the ACE interviewing sample.

This subsampling operation will reduce the sample of small clusters while at the same time attempting to balance three goals. First, we would like to prevent any small clusters from having weights that are extremely high compared to other clusters in the sample. Second, we would like to have lower weights on clusters where the number of housing units is different than we expected. These first two goals attempt to reduce the contribution of small clusters to the variance of the dual system estimates. The third goal is to ensure that the field staff can efficiently manage the resulting workloads.

Using the keyed independent listing counts and the January DMAF count, the small block clusters within each state are stratified by size. Then, a systematic subsample is selected from each stratum with equal probability. All small block clusters which have 10 or more housing units, are List Enumerate, or are on an American Indian Reservation or other type of American Indian Country are kept in the ACE sample.

Approximately 1500 of the 5000 small block clusters are expected to be retained in the ACE sample. This is the last stage of cluster reduction; therefore, there is expected to be approximately 12,500 block clusters in the ACE sample for the 50 states and the District of Columbia. For Puerto Rico, there will be roughly 500 block clusters.

2.3 Obtain DMAF Extract

The census housing units included on the Decennial Master Address File (DMAF) in January 2000 in the block clusters after sample reduction and small block subsampling are obtained for the housing unit matching. Changes made to the housing unit inventory in the DMAF after

- Match all multi-units with alphabetic designations such as A, B, C, D, etc. to multi-units with numeric designations such as 1, 2, 3, 4, etc. In other words: A=1, B=2, C=3, D=4, E=5, ... Y=25, Z=26.

- Match all multi-units with the following set of equivalent expressions:
  1) Apt 1=Apt A=Downstairs=Right=Lower=Front
  2) Apt 2=Apt B=Upstairs=Left=Upper=Rear

The results of the computer matching and additional information are loaded into a clerical matching support database which is utilized by the clerical matchers. This database is called the Matching Review and Coding System (MaRCS). The addresses will be in order by ACE map spot number. In multi-unit structures, the individual apartments will be in within map spot ID order. The census housing units that do not match to the ACE are in census ID order within block after the last ACE housing unit.

2.4.2 Clerical Match

The clerical matchers use the results of the computer matching to aid in their matching of addresses from the ACE and the census. Only clusters expected to benefit from clerical matching are sent for clerical matching.[4] Supplemental materials are provided to facilitate the matching, such as the map spotted ACE and census maps in the rural site. Before follow-up clerical match codes are assigned for each ACE and census address that could not be matched by the computer. The matched addresses are not targeted for review, because the quality of the matches assigned by the computer has been proven to be good. Clerks are allowed to correct any errors in the computer matching that they identify while they are attempting to match the not matched housing units.

The clerical matchers use all housing unit information available to match housing units. The urban areas are almost totally city style addresses. Any addresses for housing units that are not matched for both the ACE and the census after the clerical review are sent to the field for a follow-up interview.

In rural areas, the addresses are more difficult to match, predominantly because of the non-city style addresses. The city style address and the non-city style address are in two different locations in the census advance listing address register and on the DMAF. The matchers have householder names and location descriptions to help in matching for ACE and census addresses in rural areas. The map spotted maps for the ACE and the census in rural areas are used for the final determination that the housing units matched. Computer images of the ACE and census map spotted maps which will be used in the housing unit matching are accessed by the Map Retrieval System (MRS) in MaRCS.

---

[4] See section 2.5 for discussion on clusters sent directly to field follow-up without clerical review.

There is a separate field in the database to link possible duplicates for both the ACE and the census. In before follow-up housing unit matching, the primaries to be linked to duplicates are housing units with codes of M, NE, NI, and RV. The housing units coded DE, DI, and P are not allowable primaries.

2.4.3 Technician Review

The Technicians will perform two activities in the housing unit matching: review the difficult cases and perform the quality assurance for the clerical matching. The difficult cases are the ones that the clerical matchers could not code (i.e., RV) or ones where they see something unusual. The clerical matchers have been instructed to use a code that flags the cases for a higher review. This speeds up the matching and increases the efficiency of the matching. When an RV code is entered for a case by the clerical matcher, the case is automatically flagged for review by the Technicians. In addition to clusters with RV codes, they will review clusters with address changes or follow-up notes.

The Technicians will also be responsible for the quality assurance for the housing unit matching. All of the work done by the clerical matchers will be reviewed initially until the clerical matcher is determined to be performing at an acceptable level of quality. The number of records to be reviewed before the clerical matcher is classified as acceptable is 200. After a clerical matcher's work is acceptable, a systematic sample of clusters completed by each clerical matcher will be reviewed for quality assurance. The software will continue to assess the level of quality of the clerk's work. If the clerk's work in the sample of reviewed clusters falls below the acceptable level of quality, the clerk will return to having all of their work reviewed by technicians.

If there are address records created in ACE that were the result of keying errors, the clerks code the case with an RV and the technicians will use the ZI code to remove them from processing.

ZI = The ACE address is incorrectly included in the ACE list of housing units. This error is identified clerically after the ACE list is created and does not need to be sent to the field for an interview. This code removes the address from further processing.

The technicians will not be able to correct ACE map spot numbers for the 2000 ACE. In the 1995 and 1996 tests, the errors in map spot numbering were corrected by deleting the listing for the one with an error and adding the address with the correct map spot number. These map spot errors are made by the lister or there was a keying error. The editing and quality assurance were increased for the keying for dress rehearsal, but there still were a few errors. Housing units with incorrect map spot numbers will be removed from the ACE listing and the P-sample by entering the ZM code.

ZM = The map spot number associated with a housing unit is in error. A delete code of ZM is entered for that housing unit. This code removes the

2. Clusters must meet both additional conditions:
    a. The absolute value of the difference in NI and NE (|NI-NE|) is greater than or equal to eleven.
    b. The difference between the sum of P, NI, and NE and the absolute value of the difference in NI and NE is less than or equal to fifteen.

This criteria attempts to identify clusters where the payoff of clerical review is small. If the clerks do little in the clerical matching, we might as well send them to the field without clerical review. This way the clerical matchers can concentrate on the more difficult clusters where the review will be beneficial and will reduce follow-up. If the clerical review does not reduce follow-up, there is no reason to do it.

2.6 Housing Unit Follow-up

The cases coded as not matched, possibly matched, or possible duplicates are sent for follow-up interview for all types of basic street address codes. Selected matched cases are also sent for additional information. Specifically, the cases identified for field follow-up are:

- The ACE addresses with a before follow-up code of NI. These ACE housing units are map spotted and information is obtained to confirm if they were housing units within the block cluster.
- The census addresses with a before follow-up code of NE. Information is obtained to confirm if these housing units were housing units within the sample cluster.
- The possible matches are sent to the field to determine if the ACE and census addresses refer to the same housing unit. If they do not match, then they are interviewed as an ACE nonmatch and a census nonmatch during the housing unit follow-up.
- Census housing units that are identified as possible duplicates are followed up to determine if the two census addresses refer to the same housing unit.
- ACE housing units that are identified as possible duplicates are followed up to determine if the two ACE addresses refer to the same housing unit.
- Matched housing units with a unit status code of 2, 3, 4, 7, and 8 (2 = under construction, 3 = future construction, 4 = unfit for habitation, 7 = vacant trailer site in a mobile home park, 8 = other).

The ACE housing unit with unit status indicating something other than an occupied or vacant housing unit that is intended for occupancy need a follow-up interview to determine their status at the time of the follow-up interview.[5] The address will be either classified as a housing unit or

[5] Housing units classified as vacant and boarded up or housing units used to store household goods fit the census definition of a housing unit, but are not occupied. These units matching the census are not followed up because they fit the census definition of a housing unit.

2.7 After Follow-up Coding

2.7.1 Clerical Coding

After the field follow-up, the completed forms are returned to the processing office. Using the information obtained during the field work, an After Follow-up Match Code is assigned for cases sent to the field. These match codes are:

M	=	The ACE and census addresses match.
CI	=	The ACE housing unit existed as a housing unit at the time of the follow-up interview and is correctly geocoded in the block cluster. The housing unit is not found in the census.
CE	=	The census housing unit existed as a housing unit at the time of the follow-up interview and is correctly geocoded in the block cluster. The housing unit is not found in the ACE.
ZI	=	The ACE address did not refer to a housing unit at the time of the follow-up interview. The ACE housing unit is removed from the enhanced list. For example, the housing unit burned, the mobile home moved, the address is commercial property, or a special place at the time of the follow-up interview.
EE	=	The census housing unit is erroneously listed on the DMAF, because the address is not a housing unit at the time of the follow-up interview in the block cluster. For example, the housing unit burned or the mobile home moved. Another example, the address is commercial property or a special place. Also, the address is nonexistent within the sample cluster.
GI	=	The ACE housing unit existed as a housing unit at the time of the follow-up interview, but is incorrectly listed in the block cluster. The housing unit is an ACE geocoding error.
GE	=	The census housing unit existed as a housing unit at the time of the follow-up interview, but is incorrectly geocoded to this block cluster. This housing unit is erroneously enumerated in this block cluster, because of a geocoding error.
DE	=	The census housing unit is erroneously enumerated in the census. The reason for erroneous enumeration is the address is duplicated in the census.
DI	=	The housing unit should not have been listed in the ACE. This address is a duplicate of another ACE address. This address is removed from further processing for the ACE and is not included on the enhanced list.
MU	=	The ACE and census addresses match and there is not enough information on the follow-up form to confirm this match as a housing unit with certainty. The follow-up interview was not done, was incomplete, was never sent, had contradictory information, or was a noninterview.
UI	=	Not enough information on the follow-up form to assign a code to the

matcher's work after they are classified as qualified for clerical matching.

The Technicians will be able to use the ZM code in after follow-up coding, since some clusters will be seen for the first time by clerks.

ZM = The map spot number associated with a housing unit is in error. A delete code of ZM is entered for that housing unit. This code removes the housing unit from further processing.

2.7.3 Analyst Review

The analysts will review the clusters with the review code and perform a quality assurance for the clusters processed in the after follow-up housing unit matching by the technicians. A systematic sample of clusters will be selected for quality assurance from each technician's work after they are classified as qualified to be a technician.

2.8 Relisting for Clusters with ACE Geocoding Errors

There may be ACE geocoding errors in the original ACE housing unit listings. These housing units are coded GI during the after follow-up coding. If a large proportion of the ACE housing units in the cluster are coded GI, the cluster will be relisted. The regional office will be notified and a blank listing book will be used to list the housing units in the cluster again. The field lister must be someone who has had no contact with this cluster before. Clusters are identified for relisting when the after follow-up matching is completed. The decision to relist is automated. If 80 percent of the cluster is coded GI[7], the cluster is relisted. There were 62 relisted clusters in the fifty states and the District of Columbia and 14 in Puerto Rico.

The ACE housing units must be collected independent of the census housing units. To have independence, the ACE housing unit listings must be done without the ACE lister seeing the census inventory of housing units. If the ACE lister sees the census housing units, independence is violated. We collect the ACE housing units and match them to the census. If there exists a large number or proportion of housing units that do not match, there is a possibility of geocoding error for the ACE or census housing units. If the ACE housing units are incorrectly geocoded to this cluster, we must start over and relist the cluster for ACE using a different lister. If the follow-up interviewer corrects the ACE listings, independence is violated because the person collecting the ACE listing of housing units has seen the census.

The number of relisted clusters is dependent on the quality of the TIGER maps, ACE listing, quality assurance not detecting errors, and housing unit follow-up. It is impossible to predict the number of clusters needing relisting in 2000. The relisting may also not be uniform across regions. There is no maximum number of relisted clusters in any region.

---

[7] If GI / (GI + M + CI) is 80 percent or greater, the cluster will be relisted.

a. If the HCUF housing unit corresponds to a housing unit on the subsampled preliminary enhanced list, then it is selected for the E sample only if the corresponding housing on the subsampled preliminary enhanced list is in a segment selected for the P sample during large block cluster subsampling.

b. The HCUF housing units that do not correspond (those housing units not in the housing unit match and those that were determined to be erroneous enumerations) to housing units on the subsampled preliminary enhanced list will be assigned to one of the segments created during the large block cluster subsampling process. HCUF housing units that are eligible to be in the E sample are those housing units that are assigned to a segment selected to be in the P sample. If there are 80 or more of these HCUF housing units then a systematic subsample will be taken with the selected HCUF housing units being in the E sample. If there are fewer than 80 of these HCUF housing units, all of these housing units will be in the E sample.

3. For relisted clusters or list/enumerate clusters there will be no housing unit matching before the E-sample identification is done. For these clusters none of the HCUF housing units correspond to housing units on the subsampled preliminary enhanced list because they did not go through the housing unit matching process. For such clusters there will be a systematic sample taken when there are 80 or more housing units. In this case the P sample and E sample will not overlap. On the subsampled preliminary enhanced list there are clusters that have no census housing units. These are clusters that contain only CI and UI units; thus, the HCUF housing units cannot be assigned to segments. If there are 80 or more HCUF housing units in such clusters, then a systematic sample will be taken.

Here are some scenarios to help illustrate the above rules.

1. Consider a block cluster that is very large with 1000 housing units in each segment created during the large block cluster subsampling process. One of these segments is selected to be in the P sample. So after large block cluster subsampling there are 1000 housing units in the P sample. If there are 1000 housing units on the HCUF that correspond the selected housing units on the subsampled preliminary enhanced list, then all of these housing units will be in the E sample.

2. Consider a block cluster with 79 A.C.E. housing units and 5 census supplemental housing units on the subsampled preliminary enhanced list and 84 housing units on the HCUF that correspond to the housing units on the subsampled preliminary enhanced list. This block cluster would not be subsampled for the P sample due to the fact that there are fewer than 80 A.C.E. housing units in the cluster. All of the housing units will be assigned to one segment and this segment will be selected to be in the P sample. It also would not be subsampled for the E sample because all of the HCUF housing unit correspond to housing units on the subsampled preliminary enhanced list and all of the housing units on the subsampled preliminary enhanced list were selected for the P sample.

- Census housing units with unresolved enumeration status.

The following housing unit codes are on the preliminary enhanced list: M, MU, UI, UE, CI, and CE. Any ACE and census housing units without enough information from the follow-up interview are unresolved. These unresolved housing units are included on the preliminary enhanced list.

The objective of the follow-up interview was to create an accurate listing of all housing units in the block cluster and to link different versions of the housing unit address that could appear on the independent listing and census lists. This updated listing is used as the inventory of housing units for subsampling large blocks. This revised listing of housing units confirmed to be in the block cluster after the housing unit follow-up is called the preliminary enhanced list. The addresses in the census that were not originally listed in the ACE listing book are not eligible to be included in the P sample.

After the housing unit matching is completed, all housing units on the preliminary enhanced list are ordered by ACE map spot number. This facilitates the subsampling. Any census housing units that did not match the ACE, but have been confirmed to be housing units in the block cluster after the original listing, were inserted in their proper location by adding a letter suffix to the map spot number. The individual units in multi-unit basic street address are ordered by apartment number in the ACE listing books, but the added census housing units are inserted at the end of the list of units in the multi-unit. For example, apartment 102 is a census nonmatch and was confirmed to have existed as a housing unit. It will not be inserted between 101 and 103, which are ACE housing units. Instead it will be inserted at the end of the list of ACE multi-units, which may have been after apartment 1215.

2.9.3 The Subsampled Preliminary Enhanced List

The subsampled preliminary enhanced list contains all of the housing units in the preliminary enhanced list plus the information from large block cluster subsampling which indicates whether or not a housing unit is in the P sample. This file is used for E-sample identification and creation of the enhanced list.

2.9.4 The Enhanced List

The enhanced list is a list of all of the housing units on the subsampled preliminary enhanced list that are to be interviewed during person interviewing. The enhanced list does not contain census housing units not in the list of ACE housing units. These census housing units are coded CE or UE. In previous tests, the housing units on the subsampled preliminary enhanced list that are in the census and not in the ACE (i.e., CE and UE) were interviewed to reduce the person follow-up workload for E-sample people who do not match to the P sample. For 2000, only the P-sample housing units will be interviewed, but the subsampled preliminary enhanced list will still be created to aid in E-sample identification. Clusters with a high proportion of ACE housing units

necessary.

After the names and characteristics are obtained, the resident status on census day is established. For nonmovers and outmovers, mover status in addition to questions about group quarters and other residences on census day establish the residence status[8].

In 1995 and 1996, the housing unit follow-up interview was conducted after census day and the status of the housing unit on census day was established during the interview. The follow-up interview is scheduled to begin before census day for the 2000 ACE.Therefore, the wording of the questions was changed in dress rehearsal to establish the status of the housing unit at the time of the follow-up interview. Since the follow-up in the housing unit phase did not identify the housing units that did not exist on census day, the ACE person interviewers will identify these housing units and code them as nonexistent on census day.

The interviewers are not allowed to add housing units during the ACE person interview, because people obtained by adding housing units to the enhanced listing are not in the P-sample for Dual System Estimation. Also, the option to add housing units was used incorrectly in 1995 and 1996. Address corrections were made by coding the housing unit as nonexistent and adding the housing unit with a corrected address instead of correcting the address. The housing unit was removed in error, resulting in the housing unit being removed from the P-sample.

The interviewers are allowed to make changes to the address during the person interview. These changes should only be updates and changes that help uniquely identify the housing unit. For example, the interviewer should add house numbers when no house number was listed in the ACE address. The interviewer is not to make changes such as correcting 1101 Hollyhills Rd to 1101 Holly Hills Rd.

3.1.1 Interviewing Outcome Codes

The interviewing outcome codes are assigned within the person interview to identify complete and partial interviews, the types of noninterviews, vacant housing units, and units that did not exist as housing units.

Outcome 200 - Unopened Case - This is a new case that has not been opened or is a case that has been sent to supervisory review and has been reassigned and not opened by the new interviewer. This outcome will never leave the laptop with the initialized action code of 00. The interviewer must take some action to resolve it.

Outcome 201- Complete Interview - This is a completed interview. We have name, age and sex (no "don't know" or "refused") for each person in the current household. We have no "don't know" or "refused" answers for any of the probes for additional people or additional residences.

---

[8] See Section 6.9 for census residence rules.

If the inmover was in group quarters and is between 18 and 22, inclusive, the inmover will be given an ACE Status code of R, removed.

3.1.3 ACE Mover and Residence Status Codes

In 1995 and 1996, the CAPI instrument identified people who lived at the sample address on census day. Each ACE person was assigned a residence status code of resident or nonresident of the housing unit on census day. If the residence status could not be determined from the information collected in the interview, the residence status was unresolved.

With procedure C, each ACE person is assigned an ACE mover code, an ACE born since census day code, an ACE group quarters code, and an ACE other residence code. The ACE status code combines all of the information from these codes to identify who is sent to person matching. The flowchart in Attachment 5 describes the process for identifying the ACE status code and who is sent to person matching. Procedure C was used in the Dress Rehearsal Accuracy and Coverage Evaluation and will be used in the 2000 Accuracy and Coverage Evaluation. See Section 6.1 for more on Procedure C.

- **ACE Mover Code**
  - 1 = Nonmover
  - 2 = Inmover
  - 3 = Outmover

- **ACE Born Since Census Day Code**
  - 0 or blank = Default for inmovers
  - 1 = Not born since census day
  - 2 = Born since census day
  - D = Don't know
  - R = Refused

- **ACE Group Quarters Code**
  - 0 or blank = Default for whole household inmovers[9]
  - 1 = In group quarters on census day
  - 2 = Not in group quarters on census day
  - D = Don't know
  - R = Refused

- **ACE Other Residence Code**
  - 0 or blank = Default for whole household inmovers
  - 1 = In other residence on census day

[9] Partial household inmovers will be assigned the codes of 1, 2, D, or R during the edit for CAPI data review.

4.0 The Person Phase

The P-sample people and the census people from the HCUF are computer matched within cluster. The possible matches, P-sample nonmatches, and E-sample nonmatches are clerically reviewed using an automated matching and review system. Additional matches and possible matches are identified by the clerical staff. Duplicates on both lists are also identified clerically. People with incomplete names are identified, because they do not contain sufficient information for matching and follow-up. After the matching is completed, field follow-up is conducted for selected cases and the results of the field interview are coded in the matching database.

P-sample people are people who were identified as nonmovers or outmovers and were residents of the ACE housing unit on census day. People with unresolved residence status are included with the P-sample people to attempt to resolve their residence status during the ACE person follow-up operation. Therefore, the P-sample people must have an ACE status of N (nonmover), O (outmover), or U (unresolved residence status). The P-sample people must have been listed in a housing unit that was listed during the housing unit phase of the ACE. In addition, the ACE person interview must be a complete or partial interview.

4.1 P-Sample Identification

The P-sample housing units are identified from the ACE housing units on the subsampled preliminary enhanced list. The housing unit match codes assigned to the ACE housing units are M, MU, CI, and UI. The census housing units with codes CE and UE are not in the P-sample. Interviews for ACE are only conducted at P-sample housing units. All P-sample housing units are on the enhanced list.

The nonmovers, outmovers, and people with unresolved ACE status (i.e., people with ACE status of N, O, and U) in P-sample housing units are included in the matching for ACE and are P-sample people for matching. The nonmovers and outmovers lived in the P-sample housing unit on census day. The people with unresolved residence status are included in the P-sample. Their residence status will be resolved during the person follow-up interview.

We discovered that there are inmovers in households with a mix of nonmovers, outmover, and inmovers who are asked the residency questions. Therefore, there can be inmovers with unresolved residence status. An edit will convert all inmovers with unresolved residence status to inmover. Therefore, the P-sample people in P-sample housing units for matching includes only people with mover status of nonmover and outmover and the ACE status is nonmover, outmover, and unresolved residence status.

4.2 Preliminary P-sample Estimation Outcome Codes

Preliminary P-sample estimation outcome codes are assigned to each P-sample housing unit before the computer and clerical matching. This outcome code is the matching outcome code

- subsampling if the number of E-sample housing units is too large

Subsampling large block clusters for ACE interviewing occurs when the block cluster contains 80 or more ACE housing units. Large clusters on American Indian Reservations are not subsampled. Subsampling large block clusters is done by segmenting the block cluster into groups of contiguous housing units and selecting a sample of segments. Subsampling to generate the ACE enhanced list for a cluster occurs before the ACE interviewing begins. P-sample housing units are identified from the housing units listed for ACE. The P-sample people are identified from the ACE person interview in P-sample housing units. When an ACE address is matched to a census address, we consider that census housing unit as being on the preliminary enhanced list. Nonmatched and unresolved census housing units can also be on the preliminary enhanced list.

The housing units in the census are the housing units on the HCUF. The HCUF is used to identify E-sample housing units. The people on the HCUF[11] in these E-sample housing units are E-sample people. Each census person enumerated during the census in sample clusters is assigned an E-sample indicator and an E-sample probability code. The E-sample indicator and E-sample probability code are attached to all census person and housing unit files used in person matching.

E-sample indicator[12]

1	=	The person enumerated in the census was enumerated in a housing unit that was selected to be included in the E-sample.
2	=	The person enumerated in the census was enumerated in a housing unit that was not selected to be included in the E-sample.

It is possible for E-sample housing units in a block cluster to be sampled at different rates. Probabilities of selection will be maintained at the block cluster level and the E-sample probability code will link the housing unit and the people within the housing unit to the appropriate probability of selection on the design file[13].

---

[11] Only data defined people are on the HUCF.

[12] There is an E-sample indicator of 3, which is described in Section 4.7, Targeted Extended Search. E-sample indicators 1 and 2 are in the sample cluster and 3 is in the surrounding blocks for clusters included in the targeted extended search.

[13] See DSSD Census 2000 Procedures and Operations Memorandum Series R-4, "Accuracy and Coverage Evaluation Survey: Sample Summary File and Sample Design File Documentation", prepared by Randy ZuWallack for more information.

increase in erroneous nonmatches and erroneous matches, and these in all probability will not balance out. This is not a trivial problem–even in the case of imputations."

The census person records are reviewed by both computer and clerically to identify people with insufficient information for matching and follow-up. Only people with sufficient information for matching and follow-up are allowed to be processed in the matching and follow-up interviewing phases of the person matching. The three types of insufficient information are:

- The census people are not data defined.
- The census people are data defined, but computer coded as insufficient information for matching and follow-up.
- The census people are computer coded as sufficient information, but converted clerically to insufficient information for matching and follow-up.

The first type of census people who are not data defined are not included in the E-sample. Only data defined people are included in the E-sample. These data defined people create person records in the census.

4.4.1 Census Data Defined

The term "data defined" is a term that has been used in the past at the Census Bureau to mean that a census person record has been created. The term "Total Persons" is the total number of people counted in the census at a census ID. The term "Selected Persons" is the data defined census people at a census ID. The difference is the people who are not data defined. These people have no census person record. A whole person imputation procedure is employed to create characteristic data in the census for these people.

Two characteristics are required to be data defined, where name counts as a characteristic. Name must have at least three characters in the first and last name together. The characteristics that are included in the counting are relationship, sex, race, Hispanic origin, and either age or year of birth[16]. Census records are created on the HCUF for all data defined people. Anyone who is not data defined is a whole person imputation.

The count of census people who are whole person imputations must be identified separately from the other census people with insufficient information for matching, because it is a separate term

---

[16] Person one does not automatically have a relationship of head of household like it did in 1990 and the telephone number in item 2 on the mail return questionnaire does not count as a characteristic. The age and date of birth are examined together. If age is present, age/year of birth counts as a characteristic. If age is blank, but year of birth is present, then the age/year of birth counts as a characteristic. If age and year of birth are both blank, the age/date of birth does not count as a characteristic. The month and day of birth were used in dress rehearsal in the determination of counting the age/date of birth as a characteristic, but not in the 2000 Census.

sufficient information, the matching software will not let the clerks match to the census person until the data found on the image is entered into the system and that new data makes the person sufficient information.

Complete name is defined as:

- first name[18], middle initial, and last name
- first name and last name
- first initial, middle initial, and last name

Currently only age and year of birth are used to determine if age is present when counting characteristics to determine if the person has enough data to be data defined in the census. The ACE will use the same criteria. In other words, when the age and year of birth are both blank, month and day of birth are not considered. If this criteria changes for the determination of census data defined, the ACE will not change. Only age and year of birth will be reviewed in determining sufficient information for matching for ACE.

4.4.3 Clerical Coding of Insufficient Information for Matching and Follow-up

There are cases where the name is not blank, but should be insufficient information for matching and follow-up. Census names like Mr. Doe, Donald Duck, white female, etc. will be coded insufficient information by the clerical matchers. The computer can not recognize names that are not real or are really incomplete names.

We will review the image of the census questionnaire for census people coded as insufficient information for matching and follow-up to obtain additional data that may convert them to sufficient information for matching and follow-up. We can also allow children with first names and no last names, but with an adult that has a first and last name to be converted to sufficient information. We can also look at the roster for the long forms to get names. These updates to the names will be captured into the matching software. The software will be programmed to decide if the person has sufficient information for matching and follow-up. When the person has enough information to be sufficient information for matching and follow-up, the software will convert the match code.

4.5 P-sample Insufficient Information for Matching and Follow-up

The P-sample is reviewed by the computer to identify people who have insufficient information for matching and follow-up. The P-sample rules for sufficient information for matching and follow-up are the same as the E-sample rules. Sufficient information for matching and follow-up

---

[18] The minimum number of characters to be a name is two. This applies to the first name and to the last name. Therefore, there must be two characters in the first name and two characters in the last name.

4.6.1 Computer Matching

During computer matching, the P-sample is matched to the census. However, this matching is prioritized; first the P-sample is matched to the E-sample, then any leftover nonmatches from the P-sample are matched to the non E-sample people. The Statistical Research Division computer matcher is used. The matching occurs in two steps:

• Record Pair Ranking: The standardized names from the P-sample person and the census-side person are compared along with the person characteristics. A ranking score is assigned to each pair of people and the optimal pairs are identified.

• Determination of Match Cutoffs: The optimal pairs in the cluster are reviewed to determine the cutoffs for matches and nonmatches. All pairs above the match cutoff are identified as a match. All pairs between the match cutoff and nonmatch cutoff are identified as possible matches. All pairs below the nonmatch cutoff are nonmatched. Match cutoffs are assigned conservatively so there are virtually no false matches.

The codes assigned during the computer match are:

M	=	The P-sample and census people match.
P	=	The P-sample and census people are possible matches.
NP	=	The P-sample person with sufficient information for matching and follow-up is not matched to the census.
NE	=	The E-sample person with sufficient information for matching and follow-up is not matched to a P-sample person.
KI	=	Match not attempted for the P-sample person because the person has insufficient information for matching and follow-up. The name is blank or incomplete or the name is complete but the person has only one characteristic. This is a computer assigned code and these people are suppressed from view by the matchers.
KE	=	Match not attempted for the E-sample person. The name is blank or incomplete or the name is complete but the person has only one characteristic.
N2	=	The non group quarters, non E-sample person with sufficient information for matching and follow-up is not matched to a P-sample person.
K2	=	The non E-sample person has insufficient information for matching and follow-up.
Q2	=	The non E-sample person enumerated in group quarters with sufficient information for matching and follow-up is not matched to a P-sample person.
N3	=	The non group quarters census person in the surrounding block with sufficient information for matching and follow-up is not matched to a P-sample person.

A housing unit match code of UI is assigned to the ACE housing units in the relisted clusters and the list/enumerate clusters. The census housing units in relisted clusters and list/enumerate clusters are assigned a blank housing unit code. All census housing units added to the clusters after housing unit matching have a blank housing unit match code.

The codes assigned by the clerical matchers are[19]:

M	=	The P-sample and census people match.
P	=	The P-sample and census people are possible matches. Additional follow-up is needed.
NP	=	The P-sample person is not matched to a census person.
NE	=	The E-sample person is not matched to a P-sample person, additional follow-up is needed.
KP	=	Match not attempted for the P-sample person, because (1) the name is incomplete, such as "Mr. Jones", or (2) the name is not a valid name, such as "White Female" or "Donald Duck". This is a clerically assigned code.
KE	=	Match not attempted for the census person. The name is incomplete or not a valid name, such as "Child Jones", or "Mickey Mouse". This is a clerically assigned code.
DP	=	The P-sample person is a duplicate of another P-sample person.
DE	=	The E-sample person is a duplicate of another E-sample person. The non E-sample person in the block cluster (i.e., E-sample indicator 2) is a duplicate of the E-sample person. The E-sample person is a duplicate of a census person in a surrounding block (i.e., E-sample indicator 3).
RV	=	The match status of this P-sample or E-sample person is not clear. A review by either a technician or analyst is needed to resolve this case.
FE	=	The E-sample person is a dog, cat, or other animal that should not go to follow-up. The code is also used when there is a note on the image for the census person indicating the person died before census day. The census enumeration is fictitious.
GE	=	The E-sample person is erroneously enumerated in this block cluster, because the housing unit is a geocoding error.
NC	=	The P-sample nonmatch was found on the census roster. This person in a partial nonmatch household was not matched to the census because only name was collected in the census for this person in a large household and the census person was not data defined. No follow-up interview is necessary.
GS	=	The E-sample person is enumerated in a housing unit that exists in a surrounding block.

[19] All codes are assumed to be for a P-sample or E-sample person with sufficient information for matching and follow-up unless stated as insufficient information for matching and follow-up.

After the HCUF is pulled, a record is created to identify people enumerated in housing units and people enumerated in group quarters. The census people enumerated in group quarters are included in the matching. Census people in institutional and noninstitutional group quarters in the sample cluster have an E-sample indicator of 2 and a match code of Q2. There is no computer matching to census people enumerated in group quarters.

Any census people enumerated in group quarters and having insufficient information for matching and follow-up are assigned the K2 code by computer.

Matches between P-sample people and non E-sample people in the cluster are coded as M. The matches to census people within the cluster and subsampled out of the large blocks are identified by matches to census people with an E-sample indicator of 2.

The FE code is discussed in Section 4.6.3. The GE code is discussed in Section 4.6.5.1. The NC code is discussed in Section 4.6.5.2. The GS, GC, and GU codes are discussed in Section 4.7.2.

4.6.3 Census Images

For Census 2000, all census forms will be scanned and the subsequent information will be interpreted using OCR (Optical Character Recognition). A form is identified by the census ID (MAF ID plus the form extension) or a processing ID. Images will only be available for housing units on the January 2000 DMAF. Images will not be available for census housing units added after January 2000. An address may return more than one form, including the following: original census form, Be Counted form, a foreign language form, and/or an SEQ (Simplified Enumerator Questionnaire) form. For block clusters included in the ACE, the scanned images need to be available for the clerical matchers, technicians, and analysts to use during person matching. Be Counted forms will not be available to use for viewing images, since they do not have a census ID associated with the form when data captured.

During ACE, person matching must occur as fast as possible to allow for maximum time for follow-up for a case and subsequent estimation activities. Using our data capture system, a computer nonmatch may occur because of the unknown quality of scanned images (e.g., an incorrect capture of a name). As a result, clerical matchers could reduce the number of unresolved or nonmatch cases if they had access to the scanned census form images. Using the scanned images, the matchers could readily discern differences or problems in the responses that the scanning equipment could not recognize. For ACE, images will be captured in "real time" and written to CDs. Images will only be available for the sample clusters. No images will be available for the surrounding blocks.

The clerical matchers will review census people with insufficient information for matching and follow-up to obtain additional information that may allow them to be matched. These people are coded KE. We will not suppress any of the census people with insufficient information for matching. All review of E-sample people with insufficient information for matching and follow-

Any E-sample nonmatch that is identified as a duplicate of another E-sample person (i.e., both census people have E-sample indicators equal to 1) is assigned a code of DE. The two E-sample people are linked in the MaRCS software. The E-sample person coded DE is the duplicate and the other E-sample person is the primary. The E-sample person with the code of DE can be linked to any other E-sample person, except another E-sample person coded DE, KE, FE, and GE. For example, there is a triplicate and the codes are M, DE, and DE. The M and the first DE are linked and the M and the second DE are linked. The two people coded DE cannot be linked.

The E-sample people with E-sample indicator equal to 1 are also compared to the non E-sample census people with E-sample indicator equal to 2. When duplicates are identified between a census person with E-sample indicator equal to 1 and another census person with E-sample indicator equal to 2, the number of times that the E-sample person is duplicated with non E-sample people in the cluster is needed. The formula for the probability of erroneous enumeration, is 100 times d divided by c+d+1 percent or

$$\Pr(EE) = 100 * d / (c + d + 1) \text{ percent}$$

where

c	=	number of times the E-sample person is duplicated with another E-sample person.
d	=	number of times the E-sample person is duplicated with census people in the cluster with E-sample indicator of 2.

This assumes the E-sample person has been coded as correctly enumerated. If the E-sample person is coded unresolved, the final probability of erroneous enumeration includes an imputation for unresolved enumeration status. If the E-sample person is assigned a match code that indicates erroneous enumeration, the number of times that the E-sample person is duplicated with non E-sample people is irrelevant and ignored. A person can not have a probability of erroneous enumeration that is larger than 100 percent.

We simplified the clerical coding by assigning all duplicates with the DE code. Then using the linking of duplicates between E-sample indicators of 1 and 2, the probability could be calculated. Therefore, all duplicates are coded DE for duplicates between two E-sample people and between an E-sample person and a non E-sample person. All duplicates are linked. When the E-sample person is duplicated with a non E-sample person in the sample block, the non E-sample person is assigned a code of DE. MaRCS calculates the number of times duplicated. The first time an E-sample person is duplicated with a non E-sample person, the number of times duplicated is calculated as 1. The second time the E-sample person is duplicated with a different non E-sample person in the sample cluster, the number of times duplicated is changed to 2 for the E-sample person.

No duplicate search is conducted between the non E-sample people. Therefore, there can be no

defined. Name did not count toward data definition in 1990, because all names were not captured. Two characteristics were required in 1990 to help guard against "persons" being erroneously created by the scanning capture system due to a single mark in the 100 percent area of the questionnaire.

In 1995 and 1996, a person needed one characteristic to be data defined, since the data from the test census questionnaires were keyed. Name counted in 1995 and 1996 as a characteristic. Therefore, a person record could be created when only the name was present on the census questionnaire.

The census data will be captured by a scanning technique known as imaging. Optical character recognition (OCR) will be used to interpret the names with keying from image used as a backup if the names cannot be read with sufficient confidence by OCR. Person records are created for data defined census people when there are two characteristics captured from the census questionnaire with name counting as a characteristic. Whole person imputation will still be used for housing units whose determined size is larger than its data-defined person count. This count of imputed people is a term in the dual system estimator for ACE.

The mail return short form has a continuation roster to collect names for persons seven through twelve. The mail return long form has a roster for the names of persons one through twelve. Data were collected for the first six people in the household for both long and short forms. If the large household follow-up is unsuccessful, there will be only names for persons seven through twelve for the short form and only names in the household roster for persons seven through twelve for the long form. Person records will not be created for the people in large households with only names, since they are not data defined.

The ACE will not be creating census person records to use in the E-sample for person matching, because the count of whole person imputations in the census is a term in the dual system estimator. In addition, the long form roster will not be used in large households for persons one through six to create additional E-sample people. Therefore, we will not be making any changes to the count of whole person imputations used in the dual system estimator by including the rosters for the long and short forms in the ACE process.

The data defined persons are further reviewed for ACE to identify people without complete names. Census people who do not have complete names are coded as having insufficient information for matching and follow-up. Some people may complete the household roster and leave the name blank on the person records. We will use the household roster on the long form to update the name fields for data defined census people with blank or incomplete names.

4.7 Targeted Extended Search

The targeted extended search for 2000 ACE is a two stage process. First, clusters are identified that will benefit most from expanding the search area to surrounding blocks. Second, blocks

in the first ring of blocks if the block touched the cluster of sample blocks at one or more points. This includes the blocks that touch the corner of a sample block.

The search area is expanded for these clusters for the P-sample whole household nonmatches with no address match and for the E-sample people in housing units coded as geocoding errors. This extended search is targeted at the clusters most likely to benefit from expanding the search area. In addition, the work is targeted to blocks within the search area where the geocoding error is located. In 1990, errors were made because matching and duplicate search in the surrounding blocks was not a common event. There were many housing units in surrounding blocks to search for matching people and coding census duplicates. There was anecdotal evidence of clerks who did not bother to look in surrounding blocks because they rarely found anything. We are targeting the expanded searching to eliminate clerical errors. The clerks know there is a good chance of finding matches or duplicates when we expand the search area when there is a possibility of geocoding error and we limit the search in the expanded search area to blocks containing geocoding errors.

All people in surrounding blocks are loaded into the matching software for list/enumerate clusters, but only those clusters selected for targeted extended search will be processed. We will look at the computer match rates for the list/enumerate clusters to identify any clusters needing targeted extended search. The field materials will be prepared and the field work for the census housing units will be done by headquarters personnel. The clerical work for the list/enumerate clusters selected for targeted extended search will be delayed until the field work has been completed. Since we only anticipate a small number of list/enumerate clusters in targeted extended search, this should not slow the process. The clerks will do the P-sample matching to surrounding blocks in the same way as other clusters selected for targeted extended search. The difference in coding for the E-sample housing units is they code all people in E-sample households where all E-sample people are not matched. This is the same process as the relisted clusters.

The process is outlined below and a flowchart of the process for clusters with house number and street name addresses is attached in Attachment 3 and for clusters with rural addresses is attached in Attachment 4.

4.7.1 P-sample Matching Extended Search

The search area is expanded to search the surrounding blocks for the P-sample whole household nonmatches when the housing unit was not matched in initial housing unit matching, (i.e. the housing unit match code is UI or CI).[20] Also, P-sample housing units that matched a census housing unit in the initial housing unit match, but the census housing unit was deleted and is not on the HCUF in the cluster will be included in the P-sample extended search, if it is a whole

[20] There is no searching in surrounding blocks for partial household nonmatches or for whole household nonmatches with matching addresses.

identified as geocoding errors in the housing unit phase of ACE. A field visit is conducted in targeted extended search clusters with geocoding errors to identify the housing units that exist on the ground in the surrounding blocks. These housing units should be counted in a block in the surrounding blocks when found as actually existing in that block. This visit is conducted in the same general time frame as the ACE person interview. If the housing unit exists in the surrounding blocks, the clerks will code the E-sample person as geocoded to the surrounding blocks (i.e., GS) during the before follow-up person matching. If the housing unit exists in the sample cluster, the E-sample person is coded as not a geocoding error, because that housing unit exists in the sample cluster after all (i.e., GC). A person follow-up interview for the E-sample nonmatches coded GS and GC is needed to identify other reasons for erroneous enumeration, such as fictitious people and other residences where people should have been counted on census day. If the housing unit does not exist in the surrounding blocks or can not be located on the map sent with the case, the E-sample person is coded as GE, indicating a person who is erroneously enumerated because the housing unit is incorrectly geocoded in the block cluster.

If the field work was not done or if we can not determine if the block number entered on the form is in the block cluster or in the surrounding blocks, the unresolved code, GU is used. There is no follow-up for the unresolved cases.

This work is done for households that can be assigned the code of GE, which is whole household E-sample nonmatches. See Section 4.6.5.1. If the a follow-up was done for the household and the household was not in the E-sample, the follow-up interview for targeted extended search is ignored.

In relist and list/enumerate clusters there is no housing unit matching. All relist clusters will be targeted extended search clusters, and all census units identified at the time of housing unit match will have a followup interview. During the BFU match some list/enumerate clusters may be selected for targeted extended search based on a review of the matching. All census units will be followed up in these clusters.

4.7.3 E-sample Targeted Duplicate Search

A limited search for duplicated people is conducted in the targeted extended search clusters with geocoding error in the surrounding blocks. This duplicate search is to identify people duplicated because of geocoding error. This duplicate search is done first on housing units and then on people.

If the housing unit is identified as geocoded in the surrounding blocks (i.e., GS), a housing unit duplicate search is conducted. How this is done depends on whether the cluster and its surrounding blocks consist solely of urban style addresses or whether they are some or all rural style addresses. In urban areas, search in the block where the GS should have been counted. If the housing unit is duplicated, a search is conducted to identify people duplicated. The duplicate search is conducted in the block where the duplicated housing unit is located. These people are

ones that actually exist in the surrounding blocks from the ones that exist outside the expanded search area.

There is not sufficient time to conduct another interview to determine which census housing units with geocoding error really exist in the first ring of surrounding blocks. These people in added housing units coded GE from person follow-up will be treated as unresolved and the probability of correct enumeration will be imputed. The code will be converted to GU in post processing. These new cases coded GU will be treated the same as any other GU.

Deleted housing units also present a problem when the deleted housing unit was matched in the initial housing unit matching. There is a concern that the deleted unit may have been moved to a surrounding block. A cluster with matched housing units in the DMAF that were deleted before the HCUF is created have no chance of being TES clusters. Therefore, we will allow surrounding block search for these housing units in the TES P-sample matching.

4.8 Technician Review

The ACE Technicians review cases coded RV by the clerical matchers. Other clusters with a cluster review indicator, such as high weights and high rates of nonmatch and unresolved are reviewed by the Technicians as a quality assurance activity. In addition, the initial work by the clerical matchers is included in the quality assurance at 100 percent. The amount of work reviewed before the clerical matcher is qualified to continue is not yet determined. It probably should be a number of clusters worked and should include a large cluster.

When the clerical matcher has passed the quality assurance, the later work will be selected for quality assurance at a rate of 10 percent of the clusters completed by the clerical matcher. If the clerical matcher never passes the quality assurance, the work of that clerical matcher will be reviewed at 100 percent. We will try to reassign the clerical matchers that do not qualify to do the clerical matching.

4.9 Analyst Review

The ACE Analysts review cases coded RV by the Technicians. The Technician's work will be reviewed for quality assurance in the same manner as the clerical matchers.

4.10 ACE Person Follow-up

The person follow-up is conducted to gather additional information to accurately code the residence status of the nonmatched P-sample people and the enumeration status of the E-sample people. The P-sample nonmatches do not match to the census. We want to make sure these P-sample nonmatches actually lived in the sample cluster on Census Day. The P-sample nonmatch is sent for a follow-up interview when there is a possibility the residence status is not correct, such as partial household nonmatches, whole household nonmatches when the interview was

The people enumerated in these housing units will be included in the person matching. The housing units look like any other housing unit, except the "Just in Case" box 3 is marked. The census people will not go for a follow-up interview, because these people are highly mobile and there would be a high rate of unresolved. Therefore, if there is a whole household E-sample nonmatch and the housing unit has the "Just in Case" box 3 marked, the code for the people in the household will be changed from NE to UE.

The whole and partial codes and housing unit matching codes are already being generated by computer. The conflicting household code needs to be created. When there are both P-sample whole household nonmatches and E-sample whole household nonmatches in housing units that matched, the P-sample household is conflicting.

Therefore, we will not be sending the whole household P-sample nonmatches that are P-sample interviews with a household member and have no census people (i.e., vacant or whole household imputations). In addition, we will not conduct follow-up when there is no housing unit match to the census and the P-sample interview was conducted with a household member. This will reduce the follow-up of P-sample nonmatches by 53.2 percent according to the Chicago data.

The following table summarizes which P-sample nonmatches will be sent to follow-up:

uses these reason for nonmatch codes to impute the probability of being erroneously enumerated for E-sample people with an unresolved enumeration status. The P-sample imputation of residence probability depends on whether a person is sent to follow-up. Hence, the imputation of residence probability in the P-sample is affected because the decision to send a P-sample person to follow-up is determined by the match status of the housing unit. If the matching has not been done for added census housing units, the match status may not be correct. This should not have much of an effect on either imputation or dual system estimates because of the small volume of census adds.

If the whole households are identified as not needing a follow-up interview and a person in the household has unresolved residence status, the person is sent for a follow-up interview. In other words, the unresolved residence status overrides the decision not to send the person to follow-up.

The E-sample people that are whole household nonmatches in an E-sample housing unit that was added since the housing unit matching phase, need to have their housing unit enumeration status checked during the person follow-up. We need to identify the geocoding errors, since all of the people in the household are erroneously enumerated in the cluster. If a housing unit is incorrectly geocoded, it did not exist as a housing unit in the cluster on census day. The census housing units in list/enumerate are treated as added census housing units, since the list/enumerate clusters did not get matched in the initial housing unit matching conducted in the Spring of 2000. The housing units in relisted clusters that were identified at the time of the housing unit match are not included in this geographic check, since these census housing units are all sent to targeted extended search. However, those units in relisted clusters that are added do have their housing unit enumeration checked.

In addition for the 2000 ACE we should add the following types of cases to follow-up to collect more information to verify P-sample housing unit geography. A geography section would be added to the P-sample nonmatch section with questions similar to the ones in the E-sample geography section. The geography section will be required for the following in addition to asking the person nonmatch questions:

- All P-sample whole household nonmatches in relisted clusters, clusters in list/enumerate areas, and any other clusters not included in the housing unit matching phase of the ACE.[29]
- All P-sample whole household nonmatches in clusters with a high rate of P-sample person nonmatch. High has been determine to be 45 percent.[30]

---

[29] We had large clusters in Dress Rehearsal that could not be finished in time, but we should not have this problem in 2000. As a result, all clusters except the ones relisted and in list/enumerate clusters will be included in housing unit matching.

[30] The procedures did not look at whole household nonmatches. Instead they looked at all nonmatches.

cluster on census day. More emphasis is being made in this questionnaire to obtain a good respondent before the interview in each section is complete.

The clerical preparation required in the past was error prone because it was a tedious task. It was also time consuming. Docuprint will eliminate these clerical errors and speed up the whole process.

4.10.2 ACE Person Follow-up Outcome Codes

Outcome 1 (Completed Interview):

- An interview is complete if either of the following definitions of "DONE" are true for ALL persons in a case:

1. If the respondent, answers 'yes' to Question 1- (Do you know .....?) and knows where the person was living on census day (Question 4)

Or

2. If the interviewer has successfully contacted 3 knowledgeable people, but still no one knows the person. (The person will be considered DONE and given an after follow-up code of 'Fictitious'.)

Outcome 2 (Partial Interview):

- At least 1 person, NOT ALL, is 'DONE' according to the above definitions. This person can be from any section of the questionnaire- i.e.- an ACE person, census person or possibly matched.

Outcome 3 (Refusal)

-Unable to complete interview due to a refusal by the respondent who is an occupant of the sample household and will not allow the interview to begin or will not provide the data needed.

Outcome 4 (Not at home)

Unable to complete interview due to an inability to contact a member of the household at that particular time.

Outcome 5 (Callback appointment)

Interviewer will return at a later time/date to complete the interview.

NR = The P-sample person is identified as a resident in the block cluster on census day during the ACE person follow-up interview. The P-sample person is missed in the E-sample.

NL = The P-sample person did not live at the sample address or in the block cluster on census day and was listed as a nonmover or outmover in error. This person is removed from the list of P-sample people, since he or she was collected during the person interview in error.

NN = The P-sample person is identified as a nonresident in the block cluster on census day during the ACE person follow-up interview, because the person lived in group quarters on census day or had another residence where the person should have been counted on census day according to census residence rules. This person is removed from the list of P-sample people, since he or she was collected during the person interview in error.

NU = Not enough information is collected during the ACE person follow-up interview to identify the P-sample person as a resident or nonresident in the block cluster. The residence status for the P-sample person is unresolved. This code is also used when the P-sample person is followed up to collect geographic information and that information is not collected.

KP = Match not attempted for the P-sample person, because (1) the name is incomplete, such as "Mr. Jones" or (2) the name is not a valid name, such as "White Female" or "Donald Duck".

KE = Match not attempted for the census person. The name is incomplete or not a valid name, such as "Child Jones", or "Mickey Mouse".

DP = The P-sample person is a duplicate of another P-sample person.

DE = The E-sample person is a duplicate of another E-sample person. The non E-sample person in the block cluster (i.e., E-sample indicator 2) is a duplicate of the E-sample person. The E-sample person is a duplicate of a census person in a surrounding block (i.e., E-sample indicator 3).

GP = The P-sample person is removed because the person interview was conducted at a housing unit that exists outside the sample cluster. The person follow-up identified this housing unit as a P-sample geocoding error.

GE = The E-sample person is erroneously enumerated in this block cluster, because the census housing unit is a geocoding error (i.e., counted in the block cluster in error). The E-sample person should have been enumerated elsewhere in the census. This code is only assigned to added census housing units that are determined during follow-up to be geocoded incorrectly.

MR = The ACE person follow-up interview determined that the matched person with unresolved residence status is a resident. The person is in the P-sample and is matched to a census person. If the census person is in the E-sample, the E-sample person is a correct enumeration.

MN = The ACE person follow-up interview determined that the matched person

used when the geographic work for the person followed up is not done.

The clerks, technicians, and analysts are allowed to correct before follow-up codes, when errors in the before follow-up matching are discovered during the after follow-up matching and coding. The allowable codes are KP, KE, DP, and DE. These before follow-up codes are entered into the database in the after follow-up stage of the matching. The matchers will not be doing rematching to look for errors, but there will be situations where we followed up a person who should have been insufficient information for matching, for example. Matches and duplicates may also be discovered during the after follow-up coding and can be corrected. We are not actively looking for errors, but they may be corrected when discovered.

The NC code is not used in after follow-up coding. This work should have been done in the before follow-up matching phase. There is no reason to code NC, since the code is only used to reduce the follow-up workload. Images are available in case someone wants to look at them, but the image work should have been done in the before follow-up matching phase.

Duplicate coding is not actually allowed in the batch processing. Duplicate coding is only allowed during the cluster processing by the technicians and analysts for both the P-sample and the E-sample. The clerical matchers will assign an RV code if they discover a duplicate, but this should be rare when only looking at one household at a time.

There are unresolved codes assigned to P-sample people. There are codes to indicate that a P-sample person has unresolved match status and codes to indicate that a P-sample person has unresolved residence status. The probability of being matched is imputed for the P-sample people with unresolved match status. The probability that the P-sample person is a resident is imputed when the follow-up did not give enough information to resolve the person's residence status. The probability that a P-sample person is a resident is the probability that the person is included as a P-sample person.

The N2, Q2, N3, and Q3 codes are used in the after follow-up matching and coding, but will be automatically coded by the software.

The GE code is also used for added census housing units that are discovered to exist as housing units outside the cluster. See Section 4.7.4 for more details.

Each P-sample and E-sample person will have a final code. Each P-sample person can be classified as matched, nonmatch, or unresolved match status and as a resident nonresident, or unresolved residence status. The P-sample people classified as removed are no longer in the P-sample. The E-sample people are classified as correctly enumerated, erroneously enumerated, or unresolved enumeration status. See the final match, enumeration, and residence codes in Sections 4.14, 4.15, and 4.16.

measure of the net undercount. This will give us advance warning of post-strata with potential problems before the estimation phase has been completed. If we had this in place in dress rehearsal, we would have known about the tenure problem in rural South Carolina several months early. We would have been able to start investigating the problem earlier and have documentation ready before the estimation phase is completed.

We will not restrict this work to the after follow-up person matching. This work can be done in all phases of the matching. We need the before and after follow-up data on a flow basis for housing unit and person matching. We can do things like write duplicate identification programs for the before follow-up matching to identify clusters where they did not do the duplicate search. We will start identifying problem clusters in the housing unit phase. We will look at the before follow-up housing unit and person matching results.

4.13 Additional Computer Coding

There are additional codes assigned to each P-sample and E-sample person to be used in the imputation and for analysis purposes. These codes are Whole/Partial Code, Address Code, and Basic Street Address Code. This coding is done by computer after all data are entered for the cluster in the before follow-up matching and again after the follow-up codes have been entered into the database.

4.13.1 Whole/Partial Code

1	=	Partial household nonmatch
2	=	Whole household nonmatch
3	=	Whole household match
4	=	All other

A code to identify whole household matches, whole household nonmatches, partial household nonmatches, and other is created by computer after the clerical matching is completed. Each household is assigned a code of partial household match when there is at least one nonmatch and at least one match. Each household is assigned a code of whole household nonmatch when all members of the household are nonmatches. Each person is assigned a code of whole household match when all members of the household are matches. All of the codes in the other codes column are ignored when assigning codes 1 through 3 in the whole/partial code. Each household is assigned a code of other when the above three codes do not apply. These codes are assigned for both the P-sample and the E-sample. The whole/partial code is created before follow-up and after follow-up.

4.13.3 Basic street address Code

1	=	Single unit
2	=	Multi-unit

The codes indicating single units and multi-units are assigned. The basic street address code for the census people is assigned based on the unit designation in the HCUF. If the unit designation is present, the basic street address code is a multi-unit basic street address. If the unit designation is not present, the basic street address code is a single unit structure.

4.13.4 Follow-up Flag

blank	=	No follow-up for the person and no follow-up for the household.
0	=	The household is followed up, but the person is not followed up
1	=	The person is flagged for a follow-up interview.
2	=	The clerks have written a special question for the follow-up interviewer to ask.

4.14 Final P-sample Person Match Codes

The probability of being matched is estimated for the P-sample people with unresolved match status.

4.14.1 Matched

M	=	The P-sample and the census people were matched.
MR	=	The P-sample follow-up interview determined that the matched person with unresolved residence status is a resident. The person is a P-sample person and is matched to a E-sample person.
MU	=	The ACE person follow-up interview obtained no useful information to resolve the residence status for the matched person who had a residence status of unresolved before follow-up. The P-sample person's residence status is unresolved and the E-sample person's enumeration status is unresolved.

4.14.2 Not Matched

NP	=	The P-sample person is not matched to a census person. There was no follow-up for the whole household nonmatches from person interviews with household members and the whole household nonmatches were not conflicting household nonmatches.
NC	=	The P-sample nonmatch was found on the census roster. This person in a partial nonmatch household was not matched to the census because only name was collected in the census for this person in a large household and the census person

GP = The P-sample person is removed because the person interview was conducted at a housing unit that exists outside the sample cluster. The person follow-up identified this housing unit as a P-sample geocoding error.

4.15 E-sample Person Enumeration Codes

The probability of being correctly enumerated is estimated for the E-sample people with unresolved enumeration status.

4.15.1 Correctly Enumerated

M = The P-sample and E-sample people were matched. The E-sample person is correctly enumerated.

CE = The E-sample nonmatch is identified as correctly enumerated during the ACE person follow-up interview.

MR = The ACE person follow-up interview determined that the matched person with unresolved residence status is a resident. The person is a P-sample person and is matched to an E-sample person.

4.15.2 Erroneously Enumerated[31]

GE = The E-sample person is erroneously enumerated in this block cluster, because the census housing unit is a geocoding error (i.e., counted in the block cluster in error). The E-sample person should have been enumerated elsewhere in the census.

EE = The E-sample nonmatch is identified during the person follow-up interview as erroneously enumerated.

FE = The E-sample nonmatch is determined to be fictitious in this block cluster during the follow-up interview. The person may have existed, but should not have been enumerated in the census within this block cluster. The E-sample person is erroneously enumerated in the census in this block cluster.

DE = The E-sample person is a duplicate of another E-sample person. The code is also used when the E-sample person is a duplicate of a census person in a surrounding block. The people in the E-sample housing unit are erroneously enumerated because they were counted accurately in the surrounding block and duplicated in the sample cluster.

MN = The ACE person follow-up interview determined that the matched person with unresolved residence status is not a resident in this housing unit or in this block

---

[31] The E-sample people who are duplicated with census people with E-sample indicator of 2 are not full erroneous enumerations. See Section 4.6.4 for the probability of erroneous enumeration when there is duplication between a census person in the E-sample and a census person record with E-sample indicator 2.

the P-sample. Housing units with preliminary estimation outcome codes of 6, 9, 11, or 12 have the same final estimation outcome code. Households with a preliminary outcome code of 10 have the final outcome code of 10. The outcome code may change when people are removed from the P-sample for preliminary estimation outcome codes of 1, 2, or 3. The types of cases where the final outcome codes are changed are:

- Whole households coded NN are converted to vacant (i.e., outcome code 10).
- Whole households coded FP or DP are converted to noninterview (i.e., outcome code 4).
- Whole households coded NL and MN are converted to noninterview (i.e., outcome code 4).

- Households with at least one person coded GP are converted to nonexistent (i.e., outcome code 12).

- Households with a mixture of NN, NL, FP, DP and MN are converted to noninterview (i.e., outcome code 4)

- Households with a mixture of KP, KI, NN, NL, FP, DP, MN are treated as whole households with insufficient information for matching and follow-up (i.e., outcome code 9).

The final estimation outcome codes for Census Day are:

1	=	Complete interview with a household respondent
2	=	Complete interview with a proxy respondent
3	=	Partial interview (i.e., some, but not all P-sample people have sufficient information for matching and follow-up)
4	=	No Census Day Residents - Households converted to noninterviews
6	=	Field noninterview
9	=	No people have sufficient information for matching and follow-up
10	=	No census day residents - Vacant
11	=	Vacant on census day
12	=	Not a housing unit on census day

P-sample estimation outcome codes 1 through 3 are interviews, 4, 6, and 9 are noninterviews, 10 and 11 are vacant, and 12 is not a housing unit within the block cluster.

4.17.2 P-sample Interview Status Codes for Census Day

The P-sample interview status codes for census day are defined as follows:

1	=	Interview
2	=	Noninterview

4.16.1 Resident

M = The P-sample and the E-sample people were matched.

MR = The P-sample follow-up interview determined that the matched person with unresolved residence status is a resident. The person is a P-sample person and is matched to a E-sample person.

NR = The P-sample person is identified as a resident in the block cluster on census day during the ACE person follow-up interview. The P-sample person is missed in the census.

NP = The P-sample person is not matched to an E-sample person. There was no follow-up for the whole household nonmatches from person interviews with household members and the whole household nonmatches were not conflicting household nonmatches. These people are considered residents of the housing unit on census day.

NC = The P-sample nonmatch was found on the census roster. This person in a partial nonmatch household was not matched to the census because only name was collected in the census for this person in a large household and the census person was not data defined. No follow-up interview is necessary.

4.16.2 Nonresident

FP = The P-sample person is fictitious in this block cluster. The person was interviewed in error during the person interview. This person is not included in the final P-sample.

NL = The P-sample person did not live at the sample address or in the block cluster on census day and was listed as a nonmover or outmover in error. This person is removed from the list of P-sample people, since he or she was collected during the person interview in error.

NN = The P-sample person is identified as a nonresident in the block cluster on census day during the ACE person follow-up interview, because the person lived in group quarters on census day or had another residence where the person should have been counted on census day according to census residence rules. This person is removed from the list of P-sample people, since he or she was collected during the person interview in error.

DP = The P-sample person is a duplicate of another P-sample person.

MN = The ACE person follow-up interview determined that the matched person with unresolved residence status is not a resident in this housing unit or in this block cluster. The person is no longer in the list of P-sample people and the E-sample person is an erroneous enumeration.

GP = The P-sample person is removed because the person interview was conducted at a housing unit that exists outside the sample cluster. The person follow-up identified this housing unit as a P-sample geocoding error.

4.17.4 P-sample Interview Status Codes for Interview Day

The P-sample interview status codes for interview day are defined as follows:

1	=	Interview
2	=	Noninterview
3	=	Not an occupied housing unit on interview day

If the final estimation outcome code for interview day is 1, 2, or 3, the P-sample interview status code for interview day is 1. If the final estimation outcome code for interview day is 5, 7, 8, 9, or 10, the P-sample interview status code for interview day is 2. If the final estimation outcome code for interview day is 11 or 12, the P-sample interview status code for interview day is 3.

5.0 Final Housing Unit Match

The initial housing unit matching described in Section 3 is conducted before the inventories of census housing units and the A. C. E housing units are final. To produce housing unit coverage estimates it will be necessary to perform a Final Housing Unit Match that incorporates the final state of the E-sample and P-sample housing units. The computer and clerical work for the final housing unit match will be done after the person matching is completed. The results of the final housing unit match are used to calculate housing unit coverage estimates, to weight the long form estimates and for evaluation efforts.

The Housing Unit Match is to be conducted between the ACE address listing and the January 2000 Census Housing Unit inventory. Only the updates to the Census and ACE housing unit inventories that arise between the time of the Housing Unit Match and the definition of the P-sample and E-sample will be processed during Final Housing Unit Match. In addition, housing unit matching will be conducted for the first time in relisted clusters and list/enumerate clusters during Final Housing Unit Match.

There is no computer matching for the final housing unit matching. However, there is a computer processing stage that determines what goes to the clerical stage. After the clerical matching, a follow-up interview will be conducted for selected cases that were not followed up during the initial housing unit matching.

Because of concerns about the availability of maps for matching in rural areas, the results of the person match will be available as a resource in FHU match. A household of census people that matches a household of A.C.E. people is evidence that the housing units in which the households are match.

Note that with the current ACE schedule, we may not have many of the updates from the update/leave areas included in the DMAF when the initial housing unit matching is conducted. Therefore, the workload for the final housing unit match in update/leave areas may be larger than

units instead of rematching everything. The highest ranking code[35] from the initial housing unit match is loaded into the database.

The following cases will be flagged for clerical review:

- New Census nonmatches, both E-sample and non E-sample.
- New P-sample nonmatches.
- Duplicate links between E-sample housing units coded DE and non E-sample housing units in all clusters.
- P-sample units nonmatched in initial Housing Unit match in targeted extended search clusters.
- E-sample units coded GE in targeted extended search clusters.

The units in list/enumerate and relist clusters are all treated like new nonmatches. Additional matches and possible matches are identified by the clerical matchers, with matches coded M and possible matches P. A duplicate search is conducted to identify census duplicates created by the added census housing units. If an E-sample unit coded DE is duplicated to an non E-sample unit, the clerk will code the non E-sample unit DE and the E-sample unit an appropriate code other than DE, such as M or CE.

The FHU BFU match begins by attempting to match the new non-matches. This is followed by the TES matching. The last part of BFU is a duplicate search that is performed for all the new A.C.E. and census non-matches.

The FHU match will require some follow-up, because some of the housing units have never had a housing unit follow-up interview. For example, a housing unit added to the census since the initial housing unit match may not be matched to an ACE unit. This new census housing unit needs a follow-up interview to determine if it was a housing unit in the cluster on census day. The census add could be a geocoding error. When the census add is matched to an ACE nonmatch, no follow-up is necessary. A census delete may have been a match and follow-up is needed for the new ACE nonmatch.

New E-sample nonmatches and new E-sample duplicates, new possible matches, new matches with unresolved housing unit status, new P-sample duplicates and some P-sample nonmatches that were not sent to follow-up in the initial housing unit matching are sent for a follow-up interview for the final housing unit phase of the ACE. (Most new P-sample nonmatches will not require followup because they can be coded based on the final estimation outcome code.) These cases are coded NE, DE, P, MU, DI and NI during the final housing unit matching. The follow-

---

[35]The highest ranking code from the MaRCS software is the last code given to a housing unit. If the computer matched two housing units, no codes are assigned in the later stages of matching. In other words, the codes assigned at all levels of matching are combined into one final highest ranking code.

5.4 Final Housing Unit Match Codes after Computer Processing

After the computer processing there will be the following codes:

NI	=	The P-sample address is a new nonmatch to a census address.
NE	=	The census address in the E-sample is a new nonmatch to an ACE address.
N2	=	The census address not in the E-sample is a new non-match.
RV	=	The census address is either a primary not in the E-sample or a duplicate whose primary not in the E-sample

In addition, the housing units with the following codes from initial housing unit match remain on the files: M, CI, CE, EE, GE, DE, MU, UI and UE.Note that housing units with the following codes from the initial housing unit match are removed from final housing unit match: GI, DI, ZI, ZM and ZE. See Section 2.7.1 for definitions of these codes.

5.5 Before Follow-up Match Codes

The codes resulting from the BFU processing of the new nonmatches in computer processing, NI, NE and N2, are M, P, NI, NE, DI, DE, CI, and MU. In TES clusters there are also the codes GS, GC, GU, and GE. For units not processed in final housing unit match, the match codes from original housing unit match still apply.

5.6 Final P-sample Housing Unit Match Codes

5.6.1 Matched

M = The P-sample and census addresses are matches.

5.6.2 Not Matched

CI = The P-sample housing unit existed as a housing unit at the time of the follow-up interview and is correctly geocoded in the block cluster. The housing unit is not found in the census.

5.6.3 Unresolved Housing Unit Status

MU = The P-sample and census addresses match and there is not enough information on the follow-up form to confirm this match as a housing unit with certainty. The follow-up interview was not done, was incomplete, was never sent, had contradictory information, or was a noninterview. The match status is matched, but the housing unit status is unresolved.

UI = Not enough information on the follow-up form to assign a code to the nonmatched P-sample housing unit with certainty. The follow-up

block cluster.

5.7.2 Erroneously Enumerated[36]

EE = The E-sample housing unit is erroneously listed on the HCUF, because the address is not a housing unit on census day in the block cluster. For example, the housing unit burned or the mobile home moved. Another example, the address is commercial property or a special place. Also, the address is nonexistent within the sample cluster.

DE = The census housing unit is erroneously enumerated in the census. The reason for the erroneous enumeration is the address is duplicated in the census.

GE = The E-sample housing unit existed as a housing unit at the time of the follow-up interview, but is incorrectly geocoded to this block cluster. This housing unit is erroneously enumerated in this block cluster, because of a geocoding error.

5.7.3 Unresolved

MU = The ACE and census addresses match but there is not enough information on the follow-up form to confirm this match as a housing unit with certainty. The follow-up interview was not done, was incomplete, was never sent, had contradictory information, or was a noninterview.

UE = Not enough information on the follow-up form to assign a code to the E-sample nonmatched housing unit with certainty. The follow-up interview was not done, was incomplete, was never sent, had contradictory information, or was a noninterview.

P = The E-sample housing unit is a possible match to the P-sample housing unit, but the follow-up interview was inconclusive or incomplete.

GU = The E-sample housing unit is not located in the ACE cluster, but there is not enough information to determine whether it was in the surrounding block. The targeted extended search field follow-up interview was not done, was incomplete, was never sent, had contradictory information, or was a noninterview.

---

[36] The E-sample people who are duplicated with census people with E-sample indicator of 2 are not full erroneous enumerations. See Section 4.6.4 for the probability of erroneous enumeration when there is duplication between a census person in the E-sample and a census person record with E-sample indicator 2.

proxy who is now living there. It is difficult to determine who lived at the sample address for outmovers. The present occupants do not always know who lived at the address before they moved in. The present occupants are not always knowledgeable about the date the previous tenants moved out of the housing unit. We must rely on neighbors to provide the names, demographic data, and responses to census residence questions. The rate of noninterviews for Procedure A will likely be higher than for Procedure B. Therefore, the noninterview adjustment must estimate the number of outmovers and the demographic data for these outmovers when the interview is not successful.

Procedure C uses the best features of Procedures A and B. The nonmovers and the outmovers are matched to the census enumerations within the sample cluster. The match rate for movers is estimated by matching the outmovers. The number of movers and their characteristics are estimated from the inmovers.

6.2 T-Night Enumerations

During the T-night operation, census questionnaires are distributed to places like marinas, carnivals, racetracks, campgrounds, RV parks, etc. If the people in these places have no usual residence, the people are enumerated on enumerator census questionnaires and a housing unit record is created for these people in the DMAF. If they indicate they have a usual home elsewhere, the census questionnaires are processed in the primary selection algorithm and the people are enumerated at their usual home.

The people enumerated in these housing units will be included in the person matching. The housing units look like any other housing unit, except the "Just in Case" box 3 is marked. The census people will not go for a follow-up interview, because these people are highly mobile and there would be a high rate of unresolved. Therefore, if there is a whole household E-sample nonmatch and the housing unit has the "Just in Case" box 3 marked, the code for the people in the household will be changed from NE to UE.

6.3 List/Enumerate Areas

The listings of census housing units in list/enumerate areas will not be available in time for them to be processed in the initial housing unit matching phase of ACE 2000. In order to keep the ACE on the same interview schedule as the urban and update/leave areas of the country in Census 2000, there will be no initial housing unit matching in list/enumerated areas. The enhanced list of addresses for person interviewing, which is the P-sample, will be the ACE housing units listed in the ACE listing books. Certain classes of structures may be excluded, such as unfit for habitation, but we will include all types of structures listed in the address listing for ACE in the P-sample for interviewing. There could be a change in the status of the unit between the time of listing and census day.

The ACE listing will start in September 1999 for the 2000 ACE. We are interested in the census

Listing Sample for Puerto Rico			
	Clusters	Housing Units	Housing Units per Cluster
Small	96	30	0.3
Medium	244	7,100	29.1
Large	219	43,400	198.2
Total	559	50,530	90.4

The sample is selected for small, medium, and large block clusters. Small clusters are defined to contain zero to two housing units, medium clusters contain from 3 to 79 housing units, and large clusters contain 80 or more housing units.

There will be a sample reduction to reduce the number of interviewed housing units to approximately 300,000. Block clusters in the small block cluster sample are listed. After the listing is completed, these small block clusters are subsampled. See Section 2.2.3 for more on small block subsampling. There are 420 clusters in list/enumerate areas, 10,883 in the 50 states and the District of Columbia, and 499 in Puerto Rico after sample reduction and small block subsampling. The following table lists the number of ACE housing units by original sampling stratum that were processed in before follow-up housing unit matching.

ACE Housing Units by Sampling Stratum		
	Housing Units in the 50 States and the District of Columbia	Housing Units in Puerto Rico
Small	5,483	155
Medium	184,904	9,000
Large	638,776	46,894
American Indian Reservation	9,264	0
Total	838,427	56,049

After the housing unit matching was completed, the housing units that did not exist as housing units on Census Day were removed as were the clusters needing relisting. The next table

or

$$DSE = \frac{(C - II)\left(\frac{CE}{N_e}\right)}{\frac{M}{N_p}}$$

where
DSE = the dual system estimate of the population in housing units
C = the count of people in housing units in the census
II = the count of the people in housing units who are whole person imputations, (i.e. people who are not data defined. See Section 4.4.1)
CE = the weighted estimate of the number of correct enumerations in the E-sample
N_e = the weighted number of E-sample people
M = the estimated number of P-sample matches
N_p = the estimated number of P-sample people

Next, the denominator can be expressed in terms of nonmovers and outmovers for Procedure C. This is the number of matches from nonmovers plus the match rate of outmovers times the number of inmovers over the number of nonmovers plus the number of inmovers. The match rate for the movers comes from the outmovers and the number of movers comes from the inmovers.

$$\frac{M}{N_p} = \frac{M_n + \left(\frac{M_o}{N_o}\right) * N_i}{N_n + N_i}$$

where

M_n = the weighted estimate of the number of P-sample nonmovers who matched
M_o = the weighted estimate of the number of P-sample outmovers who matched
N_o = the weighted estimate of the number of P-sample outmovers
N_i = the weighted estimate of the number of P-sample inmovers
N_n = the weighted estimate of the number of P-sample nonmovers

Balancing error occurs when inconsistent treatments are applied to the P-sample and the E-sample. For example, the P-sample search is limited to the sample block, but the E-sample work allows expanding the search area to a surrounding ring of blocks.

The 1980 Post Enumeration Program did not have overlapping search areas. The P-sample was selected from the Current Population Survey and the E-sample was selected from housing units enumerated in the census. The two samples could have had different sized search areas causing them not to balance.

Since 1980, the coverage measurement surveys conducted by the Bureau have had overlapping search areas guaranteeing balancing for the P-sample and E-sample in the search area. Under the overlapping sample design being used, the same blocks are included in the P-sample as in the E-sample. The P-sample search area is , by definition, the proper search area. The E-sample search area is chosen to be consistent with the P-sample search area.

With our current design the search area is the block cluster for most clusters. We search in one ring of surrounding blocks for only a small number of clusters. The current plan is to do targeted extended search for approximately 20 percent of the clusters. We must be certain that the searching in surrounding rings is only conducted for the clusters selected to be included in targeted extended search to guarantee a balanced design.

6.7 E-sample Indicators

An E-sample indicator identifies the location of a census enumeration within the search area.

1	=	The census enumeration is in the E-sample within the sample cluster.
2	=	The census enumeration is in the sample cluster, but was not in sample in the large cluster.
3	=	The census enumeration is in the first ring of blocks surrounding the sample cluster.

The first ring of blocks surrounding the sample cluster is defined as all blocks touching the sample cluster at one or more points.

6.8 Housing Unit Definition and Occupancy Status[37]

Living quarters are either housing units or group quarters. They are usually found in structures intended for residential use but also may be found in structures intended for nonresidential use, like commercial buildings, as well as in tents, vans, shelters for the homeless, dormitories,

---

[37] The source of this information is a memorandum from John H. Thompson to Robert W. Marx, dated February 3, 1997, Census 2000 Decision Memorandum No. 8, subject: Revised Housing Unit Definition for Census 2000.

are in place.

- Vacant units are excluded from the housing unit inventory if they are open to the elements; that is the roof, walls, windows, and/or doors no longer protect the interior from the elements. Also excluded are vacant units with a posted sign indicating that they are condemned or they are to be demolished.

- Units that are vacant, fit the definition of a housing unit, but are boarded up are housing units.

- Also excluded from the housing unit inventory are structures or rooms being used entirely for nonresidential purposes, such as a store or an office, or space used for the storage of business supplies or inventory, machinery, or agricultural products. It is a housing unit if household goods are stored in the unit.

There were two changes to the 1990 definition of a housing unit and one change to the vacancy classification. The concept of eating separately has been removed from the definition of a housing unit. Also, the conversion of housing units to group quarters when there are nine or more unrelated people living in the housing unit will not be done for Census 2000.

Vacant rooms in hotels, motels, and similar transient facilities where the building was 75 percent or more occupied by permanent residents were treated as housing units. They will not be housing units in Census 2000.

Rule 8. Person is an officer or crew member of a merchant vessel and on Census Day is engaged in inland waterway transportation.	**Count person at:** This household
Rule 9. Person works for and lives in this household and has no other home (e.g., a domestic worker or nanny who "lives in").	**Count person at:** This household
Rule 10. Person is staying temporarily in this household on Census Day and has another home.	**Count person at:** DO NOT LIST. (This person will be counted at the other household.)
Rule 11. On Census Day, person is a citizen of a foreign country who has established a household (or is part of an established household) in the U.S. while working or studying. This includes any family member living with the person.	**Count person at:** This household
Rule 12. Person is a citizen of a foreign country and on Census Day is living on the premises of an Embassy, Ministry, Legation, Chancellery, or Consulate in the U.S.	**Count person at:** This household, that is, the Embassy, etc. (The person has the right to refuse to provide any or all information.)
Group Quarters Population, UHE Allowed	
Rule 13. Person is a member of the U.S. Armed Forces and on Census Day is living on a military installation in the United States, or is living on a military vessel which is assigned to a home port in the United States.	**Count person at:** The residence where the person spends **most** of his/her time (UHE allowed) [GQ code 601 for military barracks on base; GQ code 602 for transient quarters for temporary residents; GQ code 603 for military ships]. If the person does not claim a UHE, count him/her at the military installation or at the home port of the vessel.
Rule 14. On Census Day, person is at a camp for temporary workers, such as agricultural or migrant workers; or logging, pipeline, or construction workers.	**Count person at:** The residence where the person spends **most** of his/her time (UHE allowed) [GQ code 901 for agriculture workers' dormitories on farms; GQ code 902 for other workers' dormitories]. If the person does not claim a UHE, count him/her at the camp.

Rule 21.	On Census Day, person is under formally authorized, supervised care or custody, in a nursing, convalescent, or rest home for the aged and dependent.	**Count person at:** The special place (UHE not allowed)
Rule 22.	On Census Day, person is under formally authorized, supervised care or custody, in a juvenile institution such as a residential school for delinquents.	**Count person at:** The special place (UHE not allowed)
Rule 23.	On Census Day, person is under formally authorized, supervised care or custody, in a home, school, hospital, or ward for the physically handicapped, mentally retarded, or mentally ill.	**Count person at:** The special place (UHE not allowed)
Rule 24.	On Census Day, person is at an emergency shelter, including shelters with sleeping facilities for individuals without a usual residence; shelters for abused women; shelters for runaway, neglected, or homeless children; or shelters for other homeless persons.	**Count person at:** The shelter (UHE not allowed)
Rule 25.	Person is a college student living in a group quarters (e.g., a dormitory, or sorority or fraternity house) (See Rules 5 and 6).	**Count person at:** The group quarters (UHE not allowed)
Overseas Population		
Rule 26.	Person is a member of the U.S. Armed Forces and on Census Day is stationed on a military vessel which is assigned to a home port in a foreign country.	**Count person at:** DO NOT LIST. (This person will be counted as part of the overseas population.)
Rule 27.	Person is a member of the U.S. Armed Forces and on Census Day is assigned to a military installation outside the United States. This rule includes family members living with him/her.	**Count person at:** DO NOT LIST. (This person will be counted as part of the overseas population.)
Rule 28.	Person is an American citizen overseas employed by the U.S. government and on Census Day has a place of duty abroad. This rule includes family members living with him/her.	**Count person at:** DO NOT LIST. (This person will be counted as part of the overseas population.)

Guiding Principles for the Residence Rules as They
Apply to Individual(s) with Multiple Residences

The following provides guidance for determining "usual residence" for an individual with more than one residence.

Weekly Cycle

If a person is on a "weekly cycle," he/she should be counted at the residence where he/she spends most of their time during the week. For example:

Some individuals live part of the week at a residence near where they work, and live at their "family home" the rest of the week. We consider these people to be on a "weekly cycle," and they should be counted at the residence where they spend most of their time during the week (e.g., commuter workers).

Monthly Cycle

If a person is on a "monthly cycle," he/she should be counted at the residence where he/she spends most of his/her time during the month. For example:

Some children live with one parent for one week out of the month and the other parent the remaining three weeks during the month. We consider these individuals to be on a "monthly cycle" and they should be counted at the residence where they spend most of their time during the month (e.g., children in joint custody situations).

Yearly Cycle

If a person is on a "yearly cycle," he/she should be counted at the residence where he/she spends most of his/her time during the year. For example:

1. Some individuals live in one state during the spring, summer, and fall, but move to a state in a warmer climate during the winter months ("snowbirds"). We consider these people to be on a "yearly cycle," and they should be counted at the residence where they spend most of their time during the year.

2. Some college students live at the college during the school year and at the "family home" during holidays or the summer. We consider these people to be on a "yearly cycle," and they should be counted at the residence where they spend most of their time during the year.

No Clearly Defined Cycle

If a person is on no clearly defined "cycle," he/she should be counted at the residence where

TEA 2 - Address Listing and Update/Leave

- Contains areas with some number of non-city-style (e.g., P.O. Box or Rural Route) mailing addresses
- Census address list is created from Address Listing, and updated from Address List Review (LUCA) 1999 Recanvassing (in selected areas) and Update/Leave
- Blocks are NOT included in Block Canvassing, the Postal Validation Check, or the New Construction Program
- Puerto Rico, including its military bases, is completely in TEA 2

Address Listing and Update/Leave are implemented in areas where mail often is delivered to non-city-style addresses. In these areas, it is difficult to obtain an up-to-date mailing address list and then "geocode" each address (that is, assign it to a collection block code), because of the constantly changing residential location/mailing address relationship (especially for P.O. Box addresses). The census address list therefore is compiled through a door-to-door independent listing operation (Address Listing) that is implemented in all TEA 2 blocks.

During Address Listing, enumerators knock on each residence door to obtain the occupant's name, phone number, residential address (or location description), and mailing address. (Enumerators do NOT revisit residences whose occupants are not present. This is why the census address list frequently does NOT contain a mailing address, and why the location description is the ONLY "address" in the census address list for many residences.) Enumerators identify the location of each building (containing living quarters) they encounter with a uniquely numbered map spot that they enter on their map and record in their address register; this number is linked to all residential units in the building, and stored in both the census address list and the TIGER data base. These areas will be included in Address List Review (LUCA) 1999.

At census time, enumerators deliver census questionnaires to all housing units compiled during Address Listing and that remain in TEA 2. In the course of delivering these questionnaires, the enumerators also update the census address list and map spotted map to reflect housing units that were not previously listed, and to eliminate residences that they cannot locate. (This operation is called Update/Leave, because the enumerators UPDATE the census address list and maps and LEAVE questionnaires.) Update/Leave enumerators use the residential address/location description in conjunction with the map spot location to determine the correct delivery point for all questionnaires.

Most housing units in TEA 2 areas are visited at least twice by enumerators -- once during Address Listing, and again during Update/Leave. Respondents must mail their completed census questionnaires to the Census Bureau, and so some residences also will be visited a third time, during Nonresponse Follow-up.

In some areas that otherwise meet the criteria for inclusion in TEA 2, the Census Bureau has decided that having respondents enumerate themselves and return their questionnaires via the mail is not the best way to conduct the enumeration. Some targeted populations may be less likely to return their questionnaires in the mail, and more likely to respond to an enumerator. In other areas, housing units may be vacant because they are occupied seasonally.

In these and comparable situations, enumerators visit all residences on the census address list and complete the enumeration on-site. In the course of delivering these questionnaires, they also update the census address list to 1) reflect housing units that were not previously listed (including a map spot to reflect each building that contains one or more living quarters), and 2) eliminate housing units that they cannot locate. (This operation is called "Rural" Update/Enumerate, because the enumerators work in areas that were Address Listed, UPDATE the census address list [and assign map spots as well], and ENUMERATE the residents.)

TEA 6 - Military

- Contains blocks within TEA 2 that are on military bases
- Mailout/Mailback for family housing
- Separate enumeration procedures for barracks, hospitals, etc.
- Blocks are included in both Block Canvassing and the Postal Validation Check
- These blocks are included in Address List Review (LUCA) 1998 materials, as the MAF was compiled initially in the same manner as TEA 1 areas

The Department of Defense has advised the Census Bureau that virtually all family housing (that is, individual residences as opposed to barracks, hospitals, and jails) are assigned city-style addresses to which the Postal Service delivers mail. The Census Bureau therefore implements mailout/mailback methods to enumerate the population of these individual residences. Within TEA 1 areas, blocks on military bases are assigned a TEA code of 1. Within TEA 2 areas, blocks on military bases are assigned a TEA code of 6. There is no difference between TEA 1 blocks on military bases and TEA 6 blocks in terms of either compiling the census address list or enumerating the population. Blocks within military bases in List/Enumerate areas (TEA 3) also are TEA 3.

TEA 7 - "Urban" Update/Leave

- Contains blocks initially in TEA 1
- Census address list is updated, and questionnaires are delivered concurrently, by Bureau staff (following procedures employed in TEA 2 areas, but without assigning map spots)
- Blocks ARE included in the Postal Validation Check and the New Construction Program
- The term "urban" reflects the predominance of city-style addresses, and does NOT

Some blocks that are in TEA 1 contain a significant number of living quarters with non-city-style addresses. These blocks should not be included in Block Canvassing, which is an operation that is designed to confirm and correct the existence and/or location of city-style addresses. The Geography and Field Divisions are identifying Block Canvassing assignment areas (AAs) that likely contain blocks with significant numbers of non-city-style addresses. Some of these AAs will be removed from Block Canvassing, and included in Address Listing. The blocks in these AAs will be assigned a TEA code of 9, and the census address list compilation and census enumeration activities in TEA 9 blocks will be virtually identical to those in TEA 2 blocks (for instance, they will be included in Update/Leave and Nonresponse Follow-up).

Because most of these blocks had few, if any, addresses in the MAF from the USPS, the entities the blocks are in mostly had nothing to review during Address List Review (LUCA) 1998. For this reason, most of these blocks will have their address list reviewed during a new phase of LUCA, often called "LUCA 99 ½."

6.11 Coverage Improvement Follow-up Universe

The eleven categories of cases that are eligible for Coverage Improvement Follow-up (CIFU) are:

A. Vacant or deleted housing units identified in nonresponse follow-up (NRFU)
B. Blank mail return forms
C. Lost mail return forms
D. The addresses in Panels 7, 8, and 9 of the Response Mode and Incentive Experiment that did not respond by mail or other response mode
E. Adds from New Construction operation that were not included in the NRFU Universe
F. Adds from the Update/Leave operation or the Urban Update/Leave operation that were not included in the NRFU Universe
G. Adds from the February 2000 Delivery Sequence File (DSF) that were not included in the NRFU Universe
H. Adds from the April DSF that were not included in the nonresponse follow-up universe
I. Adds from the Local Update of Census Addresses (LUCA) 98 Appeals operation that were not included in the NRFU Universe
J. Adds from the LUCA 99 Appeals operation that were not included in the NRFU Universe
K. Adds to the DMAF that were not included in the NRFU Universe and are not represented in categories E through J

For more details see DSSD CENSUS 2000 PROCEDURES AND OPERATION MEMORANDUM #CC-2 Revised #4, Definition of the Coverage Improvement Follow-up Universe for Census 2000, dated June 28, 2000.

6.12 Census Use Boxes on Short Form Nonresponse Follow-up Forms

The definitions for the census use boxes on the short form for nonresponse follow-up enumerator forms are listed below:

		Second Character	1 = Respondent filled form 2 = Enumerator interview 3 = Enumerator filled from administrative records 4 = Other
JIC4	All SEQ adds		Field Operations two digit task code

6.14 Definition II

This section is from the Hogan and Cowan paper titled Imputations, Response Errors, and Matching in Dual System Estimation" published in the 1980 Proceedings of the Section on Survey Research Methods on pages 263 to 268.

Definition I - A person is "correctly enumerated" if he should have been enumerated and was enumerated once and only once, even though it might have been in an incorrect location. A person is "missed" if he should have been enumerated in the census but was not enumerated in any location. An enumeration is considered to be an "erroneous enumeration" if the person should not have been enumerated but was (e.g., he did not exist, lived outside the U.S., was born after the census or died before the census), or the person should have been enumerated but was enumerated more than once.

Definition II - A person is "correctly enumerated" if he was enumerated in the census at the address reported by the followup survey as the census date residence. A person is "missed" if he was not enumerated at the census date residence that was reported in the followup survey. An enumeration is considered to be "erroneous" if the followup survey reports that the person was not living at the location where the census recorded him. For example, the followup survey could report that no such person exists, or that the person was born after the census, died before the census or was living elsewhere on census date.

The Census Bureau has found that it is impossible to search all locations where a person might have been enumerated. So we are forced into Definition II. But, while seemingly clear for the purpose of defining misses, the definition must be carefully used in dealing with erroneous enumerations. In theory, where one reports one should have been enumerated should be the same regardless of how one is sampled for a followup survey (System 2).[39] But for the people who move between the census and the followup survey, serious problems can arise. This brings us to our next issue: misstatement of address.

One of our serious problems is that many people misstate their Census Day address. Many people

[39] System 2 is ACE interview or the P-sample.

units. As a result, housing units or group quarters can be misclassified when they can not easily be classified as housing units or group quarters. This misclassification can be found in both the ACE and the census.

When the P-sample people in ACE housing units do not match to the census, we do match them to the people enumerated in the census in group quarters. That is, we search in group quarters for P-sample nonmatches. If the P-sample people are found in the group quarters enumerations, they are treated as matched. But, we make no attempt to discover whether the misclassification is in the ACE or the census.

Likewise, if a census person in the E-sample is enumerated in a housing unit, but the housing unit has been misclassified and should be group quarters, the follow-up of the census nonmatch will obtain information about the residence of the person. If it finds the person should be counted in this block in group quarters or a housing unit, the person will be coded as correctly enumerated in ACE processing. We do not want to classify someone as erroneous when they really should be counted in this cluster, but may have been misclassified.

If the structure and the people in it are enumerated as both a housing unit and as a group quarters, the census people who are enumerated in a housing unit are eligible to be in the E-sample. If the census person is not matched, the follow-up interview will identify them as living in the cluster and having no other residence and they will be coded as correctly enumerated. There is no duplicate search between people enumerated in group quarters and housing units.

In summary, then, we take a balanced look (1) for P-sample people not found in census housing units, we look in group quarters and (2) for E-sample people not found, we follow up throughout the whole cluster in housing units and group quarters. We have covered the population in housing units, but have no estimate of coverage in group quarters. If the housing unit was duplicated in the group quarters, we do not count the group quarters people as duplicates. Likewise, if a group quarters was missed, we do not determine the inhabitants undercounted.

7.0 Glossary

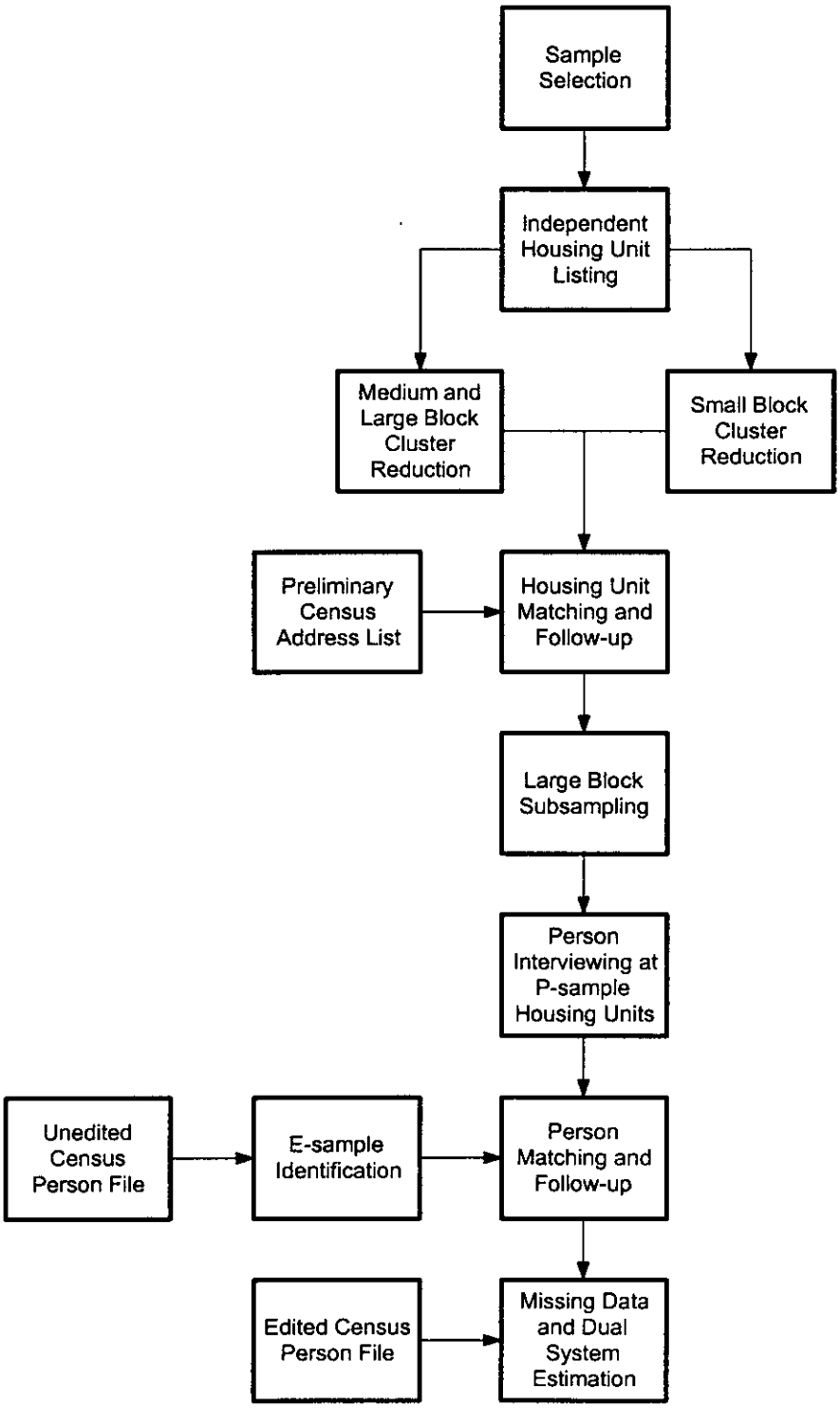
- ACE housing unit - Housing units listed in the ACE independent listing books.

- Preliminary enhanced list - The preliminary enhanced list contains all housing units confirmed to exist in the block cluster. The following housing unit codes are on the preliminary enhanced list: M, MU, UI, UE, CI, and CE. This updated listing is used as the inventory of housing units for subsampling large blocks.

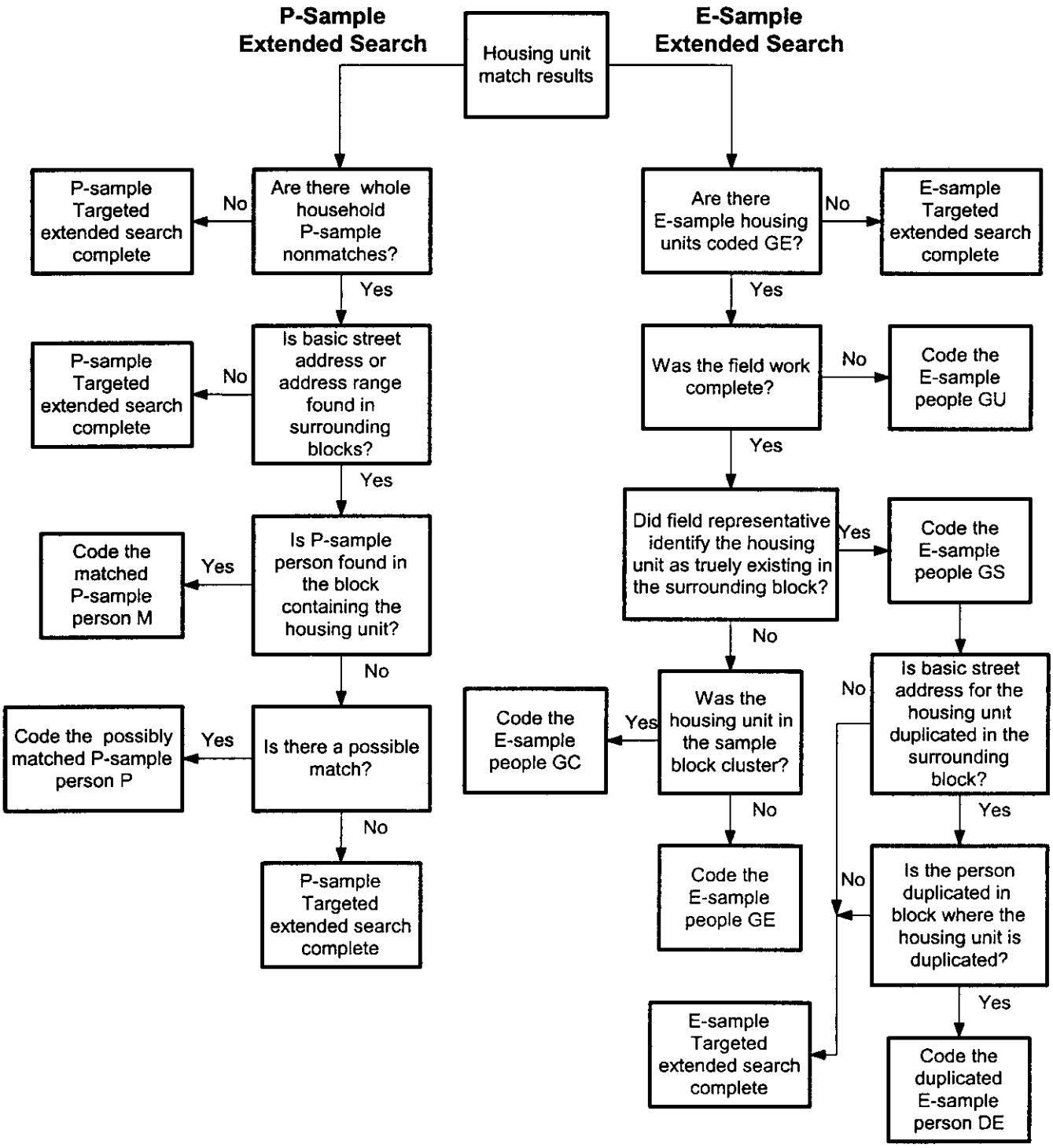
- Subsampled preliminary enhanced list - The subsampled preliminary enhanced list is the listing of housing units after subsampling is completed.

- P-sample housing unit - Housing units listed in the listing book who are confirmed to exist in

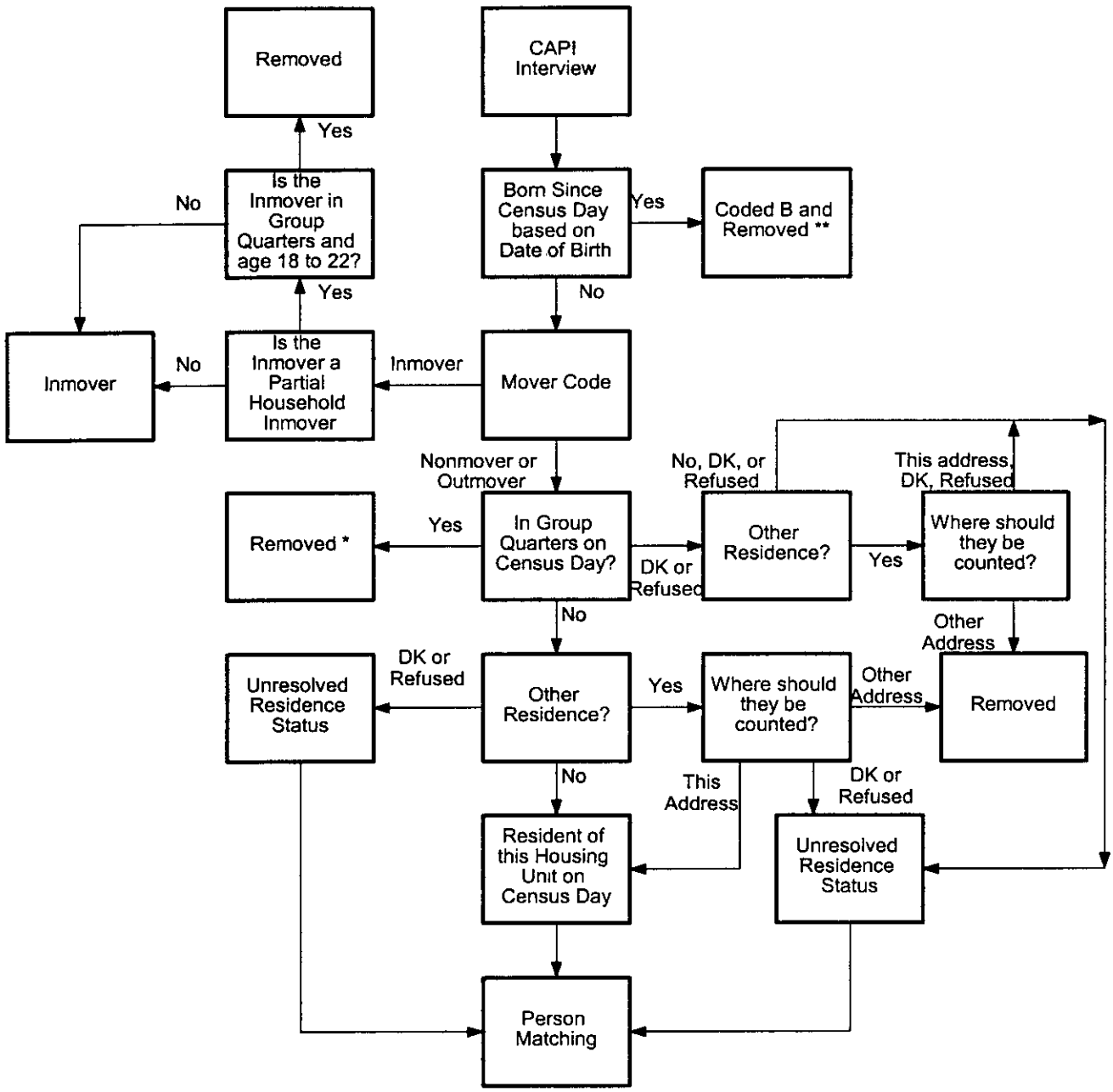
Attachment 1: Workflow for the Accuracy and Coverage Evaluation



Attachment 3
Targeted Extended Search



Attachment 5: ACE Mover and Residence Status Codes



* Other residence question is asked, but the person is removed regardless of the answer.

** Group quarters and Other residence questions are asked, but the person is removed regardless of the answer.

Conduct computer processing - final housing unit	03-07-01	04-04-01
Conduct before follow-up match - final housing unit	03-14-01	04-13-01
Conduct final housing unit field follow-up	03-30-01	05-11-01
Conduct after follow-up coding - final housing unit	04-09-01	05-22-01